Notes on Calculus

by

Dinakar Ramakrishnan
253-37 Caltech
Pasadena, CA 91125
Fall 2001

Contents

0 Logical Background **4**

0.1 Sets . . . 4
0.2 Functions . . . 5
0.3 Cardinality . . . 5
0.4 Equivalence Relations . . . 6

1 Real and Complex Numbers **8**

1.1 Desired Properties . . . 8
1.2 Natural Numbers, Well Ordering, and Induction . . . 10
1.3 Integers . . . 12
1.4 Rational Numbers . . . 13
1.5 Ordered Fields . . . 15
1.6 Real Numbers . . . 16
1.7 Absolute Value . . . 20
1.8 Complex Numbers . . . 21

2 Sequences and Series **24**

2.1 Convergence of sequences . . . 24
2.2 Cauchy's criterion . . . 28
2.3 Construction of Real Numbers revisited . . . 29
2.4 Infinite series . . . 31
2.5 Tests for Convergence . . . 33
2.6 Alternating series . . . 35

3 Basics of Integration **38**

3.1 Open, closed and compact sets in $\mathbb{R}$ . . . 38
3.2 Integrals of bounded functions . . . 41
3.3 Integrability of monotone functions . . . 44
3.4 Computation of $\int_a^b x^s dx$ . . . 45
3.5 Example of a non-integrable, bounded function . . . 47
3.6 Properties of integrals . . . 48
3.7 The integral of x^m revisited, and polynomials . . . 50

4 Continuous functions, Integrability **53**

4.1 Limits and Continuity . . . 53
4.2 Some theorems on continuous functions . . . 57
4.3 Integrability of continuous functions . . . 59
4.4 Trigonometric functions . . . 60
4.5 Functions with discontinuities . . . 64

5 Improper Integrals, Areas, Polar Coordinates, Volumes **66**
5.1 Improper Integrals . . . 66
5.2 Areas . . . 69
5.3 Polar coordinates . . . 71
5.4 Volumes . . . 73
5.5 The integral test for infinite series . . . 75

6 Differentiation, Properties, Tangents, Extrema **78**
6.1 Derivatives . . . 78
6.2 Rules of differentiation, consequences . . . 81
6.3 Proofs of the rules . . . 84
6.4 Tangents . . . 86
6.5 Extrema of differentiable functions . . . 87
6.6 The mean value theorem . . . 88

7 The Fundamental Theorems of Calculus, Methods of Integration **91**
7.1 The fundamental theorems . . . 91
7.2 The indefinite integral . . . 94
7.3 Integration by substitution . . . 94
7.4 Integration by parts . . . 97

8 Factorization of polynomials, Integration by partial fractions **100**
8.1 Long division, roots . . . 100
8.2 Factorization over $\mathbb{C}$ . . . 102
8.3 Factorization over $\mathbb{R}$ . . . 103
8.4 The partial fraction decomposition . . . 105
8.5 Integration of rational functions . . . 106

9 Inverse Functions, log, exp, arcsin, ... **110**
9.1 Inverse functions . . . 110
9.2 The natural logarithm . . . 111
9.3 The exponential function . . . 114
9.4 arcsin, arccos, arctan, et al . . . 119
9.5 A useful substitution . . . 120
9.6 Appendix: L'Hopital's Rule . . . 121

10 Taylor's theorem, Polynomial approximations **124**
10.1 Taylor polynomials . . . 124
10.2 Approximation to order n . . . 127
10.3 Taylor's Remainder Formula . . . 130
10.4 The irrationality of e . . . 135

11 Uniform convergence, Taylor series, Complex series **136**
11.1 Infinite series of functions, convergence 136
11.2 Taylor series 138
11.3 Complex power series 142

0 Logical Background

0.1 Sets

In this course we will use the term *set* to simply mean a collection of things which have a common property such as the totality of positive integers or the collection of points on a plane curve where the slope does not make sense. In truth the concept of the mathematical term *set* is a subtle one, and unless some care is exercised paradoxes will result. However, no pathology will occur in our (very) concrete situation, allowing us to be not precise about the definition of a set. Those students who are interested in learning more about these wonderful artifacts of human imagination can take Ma6c or consult the book, *Axiomatic Set theory* by Suppes.

We will often use the following symbolic abbreviations: $\exists$ for *There exists*, $\forall$ for *for all* (or *for every*), $=$ for *equals*, $\neq$ for *not equal to*, $\Rightarrow$ for *if*, $\Leftarrow$ for *only if*, $\Leftrightarrow$ or *iff* for *if and only if*, s.t. for *such that*, i.e., for *that is*, qed for *quod erat demonstrandum* (*end of proof*), $\in$ for *belongs to*, and $\notin$ for *does not belong to*.

The *members* of a set will also be called *elements*. If a set X consists of the elements $a, b, c, \ldots$, then we will write

$$X = \{a, b, c, \ldots\}.$$

Two sets will be equal *if and only if* they have the *same* elements. The empty set is denoted by the symbol $\emptyset$.

A *subset* of a set X is a subcollection Y consisting of a portion of the elements of X. We will write $Y \subset X$, or $X \supset Y$, to indicate that Y is a subset of X. Typically, Y will be given as the set of elements of X satisfying some property P, in which case we will write

$$Y = \{x \in X \,|\, x \quad \text{has} \quad \text{property} \quad P\}.$$

Clearly, the empty set is a subset of every set. We will say that Y is a *proper subset* of X if it is a subset *and* if $Y \neq X$.

If Y is a subset of X, then the *complement* of Y (in X) is the set

$$Y^c = \{x \in X \,|\, x \notin Y\}.$$

It is sometimes denoted $X - Y$.

If X, Y are two sets, their *union*, resp. *intersection*, is defined to be

$$X \cup Y = \{z \,|\, z \in X \quad \textit{or} \quad z \in Y\},$$

resp.

$$X \cap Y = \{w \,|\, w \in X \quad \textit{and} \quad w \in Y\}.$$

The following Theorem is a fundamental law, called the *de Morgan Law*:

Theorem *Let A, B be subsets of a set X. Then we have*

$$(A \cup B)^c = A^c \cap B^c$$

and

$$(A \cap B)^c = A^c \cup B^c.$$

When A, B have no intersection, we will call $A \cup B$ a *disjoint union*.

0.2 Functions

A *function*, or a *mapping*, from a set X to another, say Y, is a rule (or an assignment) f which associates, to each element x in X, a *unique* element y, denoted $f(x)$, of Y. The symbolic way of describing the function is the following:

$$f : X \to Y.$$

It is important that $y = f(x)$ be assigned uniquely to each x, but it might happen that for a fixed y, there may be many x in X with $f(x) = y$.

The *image* of such a mapping f is defined to be

$$\mathrm{Im}(f) = \{y \in Y \mid \exists\, x \in X \quad \text{s.t.} \quad y = f(x)\}.$$

Here s.t. is an abbreviation for *such that*. For any y in the image, its *pre-image*, sometimes called its *fiber*, is the set

$$f^{-1}(y) = \{x \in X \mid f(x) = y\}.$$

When $f^{-1}(y)$ is a singleton for each y in $\mathrm{Im}(f)$, we will say that the function f is *one-to-one* or *injective*. When the image of f is all of Y, we will say that f is *onto* or *surjective*.

If f is both one-to-one and onto, we will say that f is a *one-to-one correspondence* or that it is *bijective*.

When f is a function from a set X to itself, i.e., when $X = Y$, one calls f a *self-mapping*. Any bijective self-mapping of X is called a *permutation*. Show, as a simple exercise, that the number of distinct permutations of a set X with exactly n elements is n!

0.3 Cardinality

The *cardinality*, or the *order*, of a set X, denoted $|X|$, is the number of elements in it. This intuitive definition is easy enough to grasp when X is finite, i.e., when it has only a finite number of elements. But when X is *infinite*, i.e., not finite, there are various types of infinities, and we will be crude and settle on having only two types of such infinities, namely of the countable type and the uncountable type.

To be precise, let us denote by $\mathbb{N}$ the set of all counting numbers $\{1, 2, 3, 4, 5, \ldots\}$. See section 1.1 for a discussion of these numbers and their properties. We will call a set X

countable if it is either finite or if there is a one-to-one correspondence between X and $\mathbb{N}$; a countable infinite set will be said to be *countably infinite*. If X is *not* countable, we will call it *uncountable*. It will turn out that the set of all real numbers is uncountable.

Clearly, two sets have the same cardinality if there is a bijective mapping between them.

Example: *The subset* $\mathbb{N}_{\text{even}}$ *of* $\mathbb{N}$ *consisting of even counting numbers has the same cardinality as* $\mathbb{N}$*; so they are both countably infinite.* To prove it, observe that the natural mapping $f : \mathbb{N} \to \mathbb{N}_{\text{even}}$ defined by $n \to 2n$ is one-to-one and onto.

Note however, that if Y is a proper subset of a *finite* set X, we will always have $|Y| < |X|$.

A basic result on cardinality is the following

Proposition *Let* X, Y *be arbitrary finite sets. Then we have*

$$|X \cup Y| = |X| + |Y| - |X \cap Y|.$$

The proof is left as an exercise.

0.4 Equivalence Relations

Sometimes it is important to partition sets into blocks, where each block consists of elements which are equivalent in some sense. For example, we can split $\mathbb{N}$ into two blocks $\mathbb{N}_{\text{even}}$ and $\mathbb{N}_{\text{odd}}$, the former consisting of *even numbers*, and the latter consisting of *odd numbers*. The equivalence would be $a \sim b$ iff a and b have the same *parity*, i.e., iff $a - b$ is divisible by 2.

In general one has to make sure that a *relation* $\sim$ on a set X has certain natural properties in order to be called an *equivalence relation*. The properties one needs are

$$a \sim a \qquad (\textbf{reflexivity})$$

$$a \sim b \Leftrightarrow b \sim a \qquad (\textbf{symmetry})$$

$$a \sim b, b \sim c \Longrightarrow a \sim c \qquad (\textbf{transitivity})$$

Check that the relation defined above on $\mathbb{N}$ by using parity satisfies these three conditions.

Now let X be any set with an equivalence relation $\sim$. For any element x, we can put together all the elements of X which are equivalent to it, and call it the *equivalence class* of x (or represented by x). Check that the properties above preclude any element from belonging to two different equivalence classes. (Such a way of writing a set as a disjoint union of subsets is called a *partition*.)

This allows us to form a new set, often denoted $X/\sim$, whose elements are equivalence classes in X. Again, when $X = \mathbb{N}$ and $\sim$ is the relation given by parity, $X/\sim$ consist of two classes, namely $\mathbb{N}_{\text{even}}$ and $\mathbb{N}_{\text{odd}}$.

Sometimes one finds the notion of a set of equivalence classes a bit abstract and opts to do the following: Choose, for each class C, a fixed element $x(C)$ representing it in X, and view $X/\sim$ as the *set of* such *representatives* $\{x(C)\}$. In the example we have been

considering, we may choose 0, resp. 1, to represent the class of even, resp. odd numbers, and identify $\{0, 1\}$ with the set of equivalence classes. (When we say that two sets can be identified, we mean that there is a one-to-one correspondence between them.)

There is no reason to feel intimidated by taking equivalence classes. One does it often in real life without thinking about it explicitly. For example everyone is used to dealing with fractions, which are really equivalence classes, because we identify $\frac{md}{nd}$ with $\frac{m}{n}$ for any non-zero integer d. More on this in section 2.1.

1 Real and Complex Numbers

1.1 Desired Properties

Let us begin by asking what one would like to have in the number system one works with? As a child grows up, he/she first learns about the collection $\mathbb{N}$ of *counting numbers*, which mathematicians like to call the *natural numbers*, then proceeds to the all important *zero* and *negative numbers*, and after that to *fractions*, often called the *rational numbers*. What one has at that point is a very elegant system of numbers with all the desired *arithmetical operations* and the important notion of *positivity*. This last property allows one to partition the rationals into *positive numbers*, *negative numbers* and zero, and thereby define an *ordering* on the set $\mathbb{Q}$ of all rational numbers, namely the stipulation that $a \leq b$ iff $b - a$ is either positive or 0. One defines a number b to be an *upper bound* (resp. a *lower bound*) for a set X of rationals iff every x in X satisfies $x \leq b$ (resp. $b \leq x$). The set X is said to be *bounded above*, resp. *bounded below*, if it admits an upper bound, resp. lower bound; it is *bounded* if both bounds exist. Of course any finite set of rationals is bounded, but infinite sets like $\{\frac{1}{n} \,|\, n \in \mathbb{N}\}$ are bounded too.

One could stay in this rational paradise for a long time and find many treasures to unearth, but one runs into *two related problems*. The first problem is that various numbers one encounters in *real life*, for example the hypoteneuse ($= \sqrt{2}$) of any right triangle with unit sides and the circumference ($= 2\pi$) of the unit circle, are not rational numbers, though historically it took a while for people to *prove* their *irrationality*. The second problem is that given a bounded set of rational numbers there is usually no rational number b which is a *least upper bound* (*lub*), also called a *supremum*, i.e., having the property that if b_1 is any upper bound for X, then $b \leq b_1$. A natural example is furnished by the (bounded) set

$$X_0 = \{\sum_{m=1}^{n} \frac{1}{m!} \,|\, n \geq 1\}. \tag{1.1.1}$$

Similarly, a general X has no *greatest lower bound glb*, often called an *infimum*. (But X_0 does have an infimum, namely 1.)

Note that the least upper bound will be unique if it exists. Indeed, if b, b' are suprema of a bounded set X, then $b \leq b'$ and $b' \leq b$, yielding $b = b'$.

One is thus led to enlarge $\mathbb{Q}$ to get a bigger number system, to be called the set $\mathbb{R}$ of *real numbers*, which solves the second problem above while still preserving the basic properties of $\mathbb{Q}$ including positivity. We want the following properties to hold:

[$\mathcal{R}$1] $\mathbb{R}$ admits the four basic arithmetical operations;

[$\mathcal{R}$2] There is a subset P of (*positive*) real numbers, closed under addition and multiplication, such that $\mathbb{R}$ is the disjoint union of P, 0 and $-P$; and

[$\mathcal{R}$3] Every subset X which is bounded above admits a least upper bound.

In addition one also wants to make sure that this larger system is not too big; it should *hug* the rational numbers ever so tightly. To phrase it mathematically, one wants the following to hold as well:

[$\mathcal{R}4$] Given any two distinct numbers x, y in $\mathbb{R}$, there is a rational number a such that $x < a < y$.

It turns out that there is a unique number system satisfying properties $\mathcal{R}1$ through $\mathcal{R}4$, which we call $\mathbb{R}$. These properties are extremely important and will be repeatedly used, implicitly or explicitly, in Calculus and in various other mathematical disciplines.

Often one thinks of $\mathbb{R}$ as a collection of decimal expansions, with a natural equivalence (like $1.9999\cdots = 2$), and the rationals are characterized as the ones with eventually repeating blocks of digits; for example, $\frac{22}{7}$ is represented by $3.142857142857142857142857\ldots$. (A justification for this way of thinking of $\mathbb{R}$ will come later, after we learn about Cauchy sequences and completion in chapter 2.) From this point of view, one sees quickly that there is nothing terribly real abut the real numbers, as no one, and no computer, can remember the *entire* decimal expansion of an irrational number, only a suitably truncated rational approximation, like $\frac{22}{7}$, or better, $3.14159265358979323846 = \frac{314159265358979323846}{100000000000000000000}$, for π. One does encounter irrational numbers in nature, as mentioned above, but usually in *Geometry* or *Differential equations*.

In the case of the set X_0 of (1.1.1), the unique (irrational) supremum is the ubiquitous number e, whose decimal expansion is $2.71828182845904523536\ldots$.

It should be remarked that fortunately, the first problem encountered while working only with $\mathbb{Q}$ gets essentially solved in $\mathbb{R}$: The irrational numbers $\sqrt{2}$ and π, and a whole slew of others like them, belong to, i,e., represented by numbers in, the system $\mathbb{R}$. Qualitatively, however, there is a huge difference between $\sqrt{2}$ and π. The former is what one calls an *algebraic number*, i,e., one satisfying a *polynomial equation* with coefficients in $\mathbb{Q}$. A number which is *not* a root of any such polynomial is called a *transcendental number*; π is one such and e is another. For a long time people could not prove the existence of transcendental numbers, or rather they could not prove that the numbers like e and π, which they suspected to be transcendental, were indeed so. We do not know to this day the status of the numbers $e + \pi$ and $e\pi$, which is a terrific open problem. One could also show the existence of transcendental numbers by showing that algebraic numbers are *countable* while $\mathbb{R}$ is not.

Yes, $\mathbb{R}$ is great to work with for a lot of purposes. But it is not paradise either. This is because a lot of algebraic numbers are not to be found in $\mathbb{R}$. To elaborate, the irrational number $\sqrt{2}$ is a root of the polynomial $x^2 - 2$, i.e., a solution of the equation $x^2 = 2$. Why restrict to this polynomial? How about $x^3 - 2$ or $x^2 + 1$? It turns out that $x^3 - 2$, and indeed any polynomial of odd degree, has a real root (as we will see later in the course), though $\mathbb{R}$ does not contain all the roots; so things are not too bad here. On the other hand, one can prove that $\mathbb{R}$ cannot contain any root of $x^2 + 1$; the situation is the same for $x^2 + a$ for any positive number a. Historically people were so stumped by this difficulty that they called the square roots of negative numbers *imaginary numbers*, which is unfortunate. Just like the real numbers aren't so real, imaginary numbers aren't so imaginary. We will however bow to tradition and stick to these terms.

In any case, one is led to ask if one can enlarge $\mathbb{R}$ further to develop a number system solving this problem. The solution which pops out is the system $\mathbb{C}$ of *complex numbers*, which is unique in a natural sense. One has the following properties:

[$\mathcal{C}1$] The four basic arithmetical operations extend to $\mathbb{C}$; and

[$\mathcal{C}2$] Every polynomial with coefficients in $\mathbb{C}$ has a root in $\mathbb{C}$.

The property $\mathcal{C}2$ is often called the *Fundamental Theorem of Algebra*. For this reason one wants to work with $\mathbb{C}$ instead of $\mathbb{R}$. For example, this property allows one to be able to find eigenvalues and eigenvectors of real symmetric matrices, and this is important in both *Linear Algebra* (Ma1b material) and *Differential Equations* (Ma2b material).

Note however that the ordering property of $\mathbb{Q}$ and $\mathbb{R}$ (coming from their positive elements) seems to have evaporated in the context of $\mathbb{C}$. This is indeed the case, and this goes to show that everything has its drawbacks, and one has to carefully choose the system one works with, depending on the problem to be solved. In any case, much of Ma1a will deal with $\mathbb{R}$, but sometimes also with $\mathbb{C}$.

1.2 Natural Numbers, Well Ordering, and Induction

We will begin by admitting that everyone knows about the set $\mathbb{N}$ of natural numbers, namely the collection $\{1, 2, 3, \ldots, n, n+1, \ldots\}$. We are appealing here to one's intuition and not giving a mathematical definition, which is subtle. Interested students can consult Chapter I of the book *Foundations of Analysis* by E. Landau.

Given any two numbers a, b in $\mathbb{N}$, one can add them as well as multiply them, and the results are denoted respectively by $a+b$ and ab (denoted at times by $a.b$ or $a \times b$). One has

$$(F1) \qquad a + b = b + a$$

$$(F2) \qquad (a+b)+c = a+(b+c)$$

$$(F3) \qquad ab = ba$$

$$(F4) \qquad (ab)c = a(bc)$$

and

$$(F5) \qquad (a+b)c = ab + ac.$$

It is customary to call the property (F1) (resp. (F3)) *commutativity* of addition (resp. multiplication), and the property (F2) (resp. (F4)) *associativity* for addition (resp. multiplication). Property (F5) is called *distributivity.* The reason for enumerating these as (Fn), for $n \leq 5$, is that they will later become parts of the axioms for a *field.*

The fundamental property underlying the basic structure of $\mathbb{N}$ is the following

Well Ordering (WO): *Every non-empty subset S of $\mathbb{N}$ contains a smallest element m, i.e., $m \leq x$ for every x in S.*

When $S = \mathbb{N}$, m is the *unit element* 1, satisfying (for all x)

$$(F6) \qquad 1 \cdot x \;=\; x \cdot 1 \;= x.$$

A very useful thing to employ in all of Mathematics is induction, which we will now state precisely.

Principle of Mathematical Induction (PMI): *A statement P about $\mathbb{N}$ is true if*

(i) P holds for $n = 1$; and

(ii) If $n > 1$ and if P holds for all $m < n$, then P holds for n.

Lemma *WO $\implies$ PMI.*

Proof. Suppose (i), (ii) hold for some property P.
To show: P is true for all non-negative integers.
Prove by contradiction. Suppose P is false. Let S be the subset of $\mathbb{N}$ for which P is false. Since P is assumed to be false, S is non-empty. By WO, there exists a **smallest** element n of S, which is > 1 as P holds for 1. Since n is the smallest for which P is false, P must hold for *all* $m < n$. Hence by (ii), P holds for n as well. **Contradiction!** So P holds.
$\square$

Remark: In fact, PMI and WO are equivalent. Interested students are invited to try and show the reverse implication PMI $\Rightarrow$ WO.

Example: Let us prove by induction the formula

$$1^3 + 2^3 + \ldots + n^3 \;=\; \frac{n^2(n+1)^2}{4}.$$

It evidently holds for $n = 1$, and so let us assume that n is > 1, and assume by induction that the formula holds for all $m < n$. In particular it holds for $n - 1$, and we get

$$1^3 + 2^3 + \ldots + (n-1)^3 + n^3 \;=\; \frac{(n-1)^2 n^2}{4} + n^3 \;=\; \frac{n^2((n-1)^2 + 4n)}{4},$$

which yields the asserted formula as $(n-1)^2 + 4n = (n+1)^2$.

1.3 Integers

In order to get all the integers, one needs to add to $\mathbb{N}$ the number 0, sometimes called the *additive identity*, satisfying

$$(F7) \qquad a + 0 \;=\; 0 + a \;=\; a,$$

and the *negative numbers* $-a$ $(= (-a))$, a *unique* one for each a in $\mathbb{N}$, such that

$$(F8) \qquad a + (-a) \;=\; (-a) + a \;=\; 0.$$

It is customary to call $-a$ the *negative*, or the *additive inverse*, of a. One defines *subtraction* by setting

$$a - b \;=\; a + (-b).$$

One also sets

$$-0 \;=\; 0 \quad \text{and} \quad 0 + 0 \;=\; 0.$$

One can deduce various other identities from these, for example:

$$-(-a) \;=\; a, \quad \text{and} \quad a - b \;=\; -b + a \;=\; -(b - a).$$

The discovery of *zero* was a major achievement in human history, and it happened long after the appearance of natural numbers. Put

$$-\mathbb{N} \;=\; \{-n \,|\, n \in \mathbb{N}\}.$$

By the *set of integers* one means

$$\mathbb{Z} \;=\; -\mathbb{N} \cup \{0\} \cup \mathbb{N} \;=\; \{\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots\}.$$

One calls $\mathbb{N}$, resp. $-\mathbb{N}$, the *set of positive integers*, resp. the *set of negative integers*.

Note the following assertions for elements of $\mathbb{Z}$:

$(Ord1)$ Every a is either positive or negative or zero;

$(Ord2)$ If a, b are positive, then $a + b$ and ab are also positive.

This defines an *ordering* on $\mathbb{Z}$ given by

$$(Ord) \qquad a > b \Leftrightarrow a - b \in \mathbb{N}.$$

We will also put

$$a \geq b \Leftrightarrow a = b \quad \text{or} \quad a > b,$$

and

$$a \leq b \Leftrightarrow b \geq a \quad \text{and} \quad a < b \Leftrightarrow b > a.$$

We will come back to this important notion of ordering later.

Of course one can add, resp. multiply, any *finite* number of elements, say $a_1, a_2, \dots a_n$ of $\mathbb{Z}$, and we will use the symbolic notation

$$\sum_{j=1}^{n} a_j \;=, a_1 + a_2 + \ldots + a_n$$

resp.

$$\prod_{j=1}^{n} a_j \;=\; a_1 a_2 \ldots a_n.$$

We will use the convention that an *empty sum* is zero, while an *empty product* is 1.

Finally, we should note that $\mathbb{Z}$ *is countable*. Indeed, consider the function

$$f : \mathbb{Z} \to \mathbb{N},$$

defined by setting $f(0) = 1$ and for n positive,

$$f(n) = 2n \quad \text{and} \quad f(-n) = 2n + 1.$$

It is easy to check that this function is one-to-one and onto, as desired.

1.4 Rational Numbers

Given any a in $\mathbb{Z}$, we know that it has an additive inverse, namely $-a$, satisfying $(F8)$. But how about a *multiplicative inverse*? For this we need to find a number b such that

$$ab \;=\; 1.$$

But a little rumination will assure one that the only possibility is $a = b = \pm 1$. This is not good as one would like to have a multiplicative inverse for every non-zero a in $\mathbb{Z}$.

This problem is solved by the introduction of *rational numbers*. It is achieved as follows.

By an *ordered pair* of integers we will mean a pair (m, n), $m, n \in \mathbb{Z}$, where the order matters, i.e., (m, n) is considered to be different from (n, m). One denotes the collection of all ordered pairs of integers by $\mathbb{Z}^2$ (or $\mathbb{Z} \times \mathbb{Z}$). Set

$$Y \;=\; \{(m, n) \,|\, n \neq 0\}.$$

Introduce a multiplication in $\mathbb{Z}^2$ by

$$(m, n)(m', n') \;=\; (mm', nn').$$

If nn' is zero, then n or n' must be zero, and this implies that this multiplication preserves Y, i.e., the product of any two elements of Y is in Y.

Define a relation $\sim$ on Y by the requirement

$$(m,n) \sim (m',n') \Leftrightarrow mn' = m'n.$$

It is left as an easy exercise to check that that this gives an *equivalence relation* on Y.

Put

$$\mathbb{Q} = Y/\sim.$$

This is the set of *rational numbers*. As discussed in section 0.4, the elements of $\mathbb{Q}$, called the *rational numbers* (or *fractions*) are *equivalence classes* in Y. The equivalence class of any ordered pair (m,n) of integers with $n \neq 0$ is denoted $\frac{m}{n}$. By definition,

$$\frac{a}{b} = \frac{ad}{bd} \quad \forall d \in \mathbb{Z}, d \neq 0.$$

It is customary to call this the *cancellation property* of fractions. For instance,

$$\frac{2}{3} = \frac{14}{21} = \frac{100}{150} = \frac{1382}{2073}.$$

Since -1 is non-zero, we have in particular,

$$\frac{-m}{n} = \frac{m}{-n},$$

which we will simply denote by $-\frac{m}{n}$ and call it the *negative* of $\frac{m}{n}$.

Addition of rational numbers is defined by

$$\frac{m}{n} + \frac{m'}{n'} = \frac{mn' + +m'n}{nn'}.$$

The multiplication on Y gives rise to a (commutative) multiplication in $\mathbb{Q}$. Explicitly,

$$\frac{m}{n}\frac{m'}{n'} = \frac{mn}{m'n'}.$$

Guven a rational number $\frac{m}{n}$, it is customary to represent it by its *reduced fraction*, i.e., write it as $\frac{m'}{n'}$ with m', n' being relatively prime, with m', n' (necessarily) satisfying $m = dm'$ and $n = dn'$, where d denotes the *gcd* of m, n.

Note that we can view $\mathbb{Z}$ as a subset of $\mathbb{Q}$ by the identification

$$m = \frac{m}{1} \quad \forall m \in \mathbb{Z},$$

compatible with multiplication, addition and subtraction. We will not make any distinction (at all) between m and $\frac{m}{1}$.

Now we are ready to find multiplicative inverses. Suppose $a = \frac{m}{n}$ is a non-zero rational number. Then m is not zero either, and so we can consider the rational number $b = \frac{n}{m}$. Then we have

$$ab = \frac{m}{n}\frac{n}{m} = \frac{mn}{nm} = \frac{mn}{mn} = 1.$$

The last (cancellation) step is justified because mn is non-zero. Since the multiplication in $\mathbb{Q}$ is commutative, ba is also 1, and so b is the *multiplicative inverse* of a.

In particular, since $\mathbb{Z}$ is a subset of $\mathbb{Q}$, every integer m has a multiplicative inverse, namely $\frac{1}{m}$.

To repeat, we have in $\mathbb{Q}$:

$$(F9) \qquad \forall a \neq 0, \exists b \quad \text{s.t.} \quad ab = ba = 1.$$

We will say that a rational number a is *positive* if it is given by a fraction $\frac{m}{n}$ with m, n of the same sign. A rational number whose negative is positive will be said to be *negative*. This introduces an ordering in $\mathbb{Q}$ as in the case of $\mathbb{Z}$.

It will be left as an exercise to show that the properties $(F1) - (F8)$ as well as $(Ord1)$ and $(Ord2)$ hold in $\mathbb{Q}$. We have already seen that $(F9)$ holds in $\mathbb{Q}$.

It is useful to note the following

Lemma *Given rational numbers a, b with $a < b$, we can find another rational number c with $a < c < b$.*

Proof. Put $x = b - a$, which is a positive rational number. It suffices to show that there is a positive rational number y less than x, for then we can take c to be $a + y$. Write $x = \frac{m}{n}$, for some positive integers m, n. Then $y = \frac{m}{n+1}$ does the job. In fact, $\frac{m}{n+k}$ works for *every* $k \geq 1$. Done.

$\square$

Note that the proof shows that there are in fact an infinite number of rational numbers between a and b.

We conclude this section by noting that it can be shown that $\mathbb{Q}$ is a *countable* set. Try to construct a one-to-one correspondence between $\mathbb{Q}$ and $\mathbb{N}$.

1.5 Ordered Fields

Any set K admitting addition and multiplication, equipped with distinguished elements called 0 and 1, for which the properties (*axioms*) $(F1)$ through $(F9)$ hold, is called a *field*.

As we have seen above, $\mathbb{Q}$ is a field. But there are others, and here is a very simple one. Put

$$\mathbb{F}_2 = \{0, 1\},$$

and define addition by the rule

$$0 + 0 = 0, \quad 0 + 1 = 1 + 0 = 1 \quad \text{and} \quad 1 + 1 = 0,$$

and multiplication by the rule

$$0 \times 0 = 0, \quad 0 \times 1 = 1 \times 0 = 0 \quad \text{and} \quad 1 \times 1 = 1.$$

It is clear that the axioms $(F1)-(F9)$ are satisfied making $\mathbb{F}_2$ a field with only two elements. Check that there can be no field with just 1 element.

In order to do Calculus, we must restrict the fields we consider. Say that a field K is an *ordered field* if the properties $(Ord1)$ and $(Ord2)$ also hold in K relative to an ordering $\leq$. We know that $\mathbb{Q}$ is an ordered field. Is there any other? It is immediate that $\mathbb{F}_2$ is *not* an ordered field, because 1 is its own negative there.

As remarked in section 1.0, $\mathbb{Q}$ suffers in general from a lack of *least upper bounds* for its bounded subsets. For this reason we want to enlarge it. To this end, we are motivated to define a *complete ordered field* to be an ordered field K in which the following holds:

$(Cord)$ Every bounded subset X admits a least upper bound, i.e., there exists a b such that if $x \leq b' \, \forall x \in X$ for any b', then $b \leq b'$.

The first question is is to know if there is *any* complete ordered field. An affirmative answer will be given in the next section.

1.6 Real Numbers

The object of this section is to give a construction of all *real numbers* and show that they form a complete ordered field $\mathbb{R}$ containing $\mathbb{Q}$. They will satisfy the properties $(\mathcal{R}1)$ through $(\mathcal{R}4)$ of section 1.0. To some the construction might appear to be too abstract, at least on a first reading. One does not have to master the proof, but it is important to feel comfortable enough with the basic properties of real numbers and learn to use them at moment's notice. We will also discuss a different construction of $\mathbb{R}$ in the next chapter, after introducing Cauchy sequences and completion from that point of view; some will find that (*analytic*) approach more understandable, while others will like the present (*algebraic*) approach due to a nineteenth century mathematicician called R. Dedekind.

The basic idea is this. Geometrically, the rational numbers represent a special collection of points on a line L with 0 in the middle and the integers being plotted in the usual way. Intuitively, one would like to have the points on the line be in one-to-one correspondence with the real numbers. But hw to achive this when all one could define precisely are the rational numbers. Dedekind's clever observation was that giving a rational number b is the same as giving the set τ_b of all the rational numbers to the left of, i.e., strictly less than, b on L. Moreover, $b \leq c$ iff $\tau_a \leq \tau_b$, and τ_{b+c} is given by the set, to be called $\tau_a + \tau_b$, consisting of rational numbers c which can be written as $a_1 + b_1$ where a_1, a_2 are rational numbers with $a_1 < a$ and $b_1 < b$. In fact one can endow the collection $\{\tau_b\}$ with *all* the basic properties of $\mathbb{Q}$. So why not try to characterize every point on L by the set of rational numbers to its left? This idea is carried out below (in complete detail) via the device of *Dedekind cuts.*

By a *real number*, we will mean a subset ρ of $\mathbb{Q}$ such that

(i) ρ is neither empty nor all of $\mathbb{Q}$;

(ii) If a is in ρ, then ρ contains every rational number less than a;

(iii) For every a in ρ, there exists a c in ρ such that $a < c$.

A subset ρ defining a real number is often called a *Dedekind cut*, or just a *cut*.

Denote by $\mathbb{R}$ the set of all real numbers. Note that for any rational number a, the singleton set $\{a\}$ is *not* a real number as it does not satisfy (ii) or (iii) above. We have to find another way to put rational numbers inside $\mathbb{R}$.

Note that for any fixed rational number b, the set

$$\tau_b = \{a \in \mathbb{Q} \,|\, a < b\}$$

is a real number according to the definition above. This way we get a one-to-one mapping

$$\tau : \mathbb{Q} \hookrightarrow \mathbb{R}, \quad b \to \tau_b.$$

Identifying $\mathbb{Q}$ with its image, i.e., not distinguishing between b and τ_b, we may view $\mathbb{Q}$ as a subset of $\mathbb{R}$.

On the other hand, the set

$$\{a \in \mathbb{Q} \,|\, a \leq b\}$$

is *not* a real number, because property (iii) is not satisfied (for b).

There are infinitely many real numbers which are not rational. For example we have the following

Lemma *Let q be a positive rational number which is not a square in $\mathbb{Q}$. Then the set*

$$\rho_q = \{a \in \mathbb{Q} \,|\, a^2 < q \quad or \quad a < 0\}$$

defines a real number.

Proof. Clearly ρ_q is not empty or all of $\mathbb{Q}$, so (i) is satisfied. Now suppose that there exist $a \in \mathbb{Q}$ and b in ρ_q such that $a < b$. We have to show, to get (ii), that a lies in ρ_q. There is nothing to prove if a is negative or zero , so we may assume that a, and hence b, is positive and that a^2 is greater q. But then $b^2 > a^2 > q$ and we get a contradiction to the fact that b belongs to ρ_q. So a must belong to ρ_q.

It is left to prove that (iii) also holds. Pick any a in ρ_q. Since ρ_m obviously contains postive rational numbers, the assertion is clear if a is ≤ 0. So we let a be positive; write it as $\frac{m}{n}$ with $m, n > 0$. Put $t = q - a^2$, which is by definition a positive rational number. For $d \in \mathbb{N}$, put $f(d) = \frac{m^2(d+1)^2}{n^2d^2} - a^2$. Since $a^2 = \frac{m^2}{n^2}$, $f(d)$ is a positive rational number, and it suffices for us to choose d so that $f(d) < t$. But

$$f(d) = \frac{m^2(2d+1)}{n^2d^2} \quad \text{and} \quad \frac{2d+1}{d^2} < \frac{3}{d}.$$

So we need to choose d such that $\frac{3}{d}$ is less than $e := \frac{n^2t}{m^2}$. Since m, n are fixed, so is e. It comes down to picking d to be greater than $3/e$, which is certainly possible. Done.

$\square$

Since $\mathbb{Q} \subset \mathbb{R}$, $\mathbb{R}$ contains 0 and 1.

Theorem *$\mathbb{R}$ admits addition, multiplication and an ordering $\leq$, making it a complete ordered field relative to the identities 0 and 1. Thus the properties $(\mathcal{R}1)-(\mathcal{R}3)$ hold in $\mathbb{R}$. The property $(\mathcal{R}4)$ holds as well.*

Proof. The ordering on $\mathbb{R}$ is obtained by setting

$$\rho \leq \rho' \Leftrightarrow \rho \subset \rho'.$$

We will say that $\rho < \rho'$ iff ρ is a proper subset of ρ'. Clearly, if $\rho \leq \rho'$ and $\rho' \leq \rho$, then $\rho = \rho'$.

First we will prove that the property $(\mathcal{R}4)$ of section 1.0 holds in $\mathbb{R}$. Suppose ρ, ρ' are real numbers such that $\rho < \rho'$. Then by definition, there exists some rational number b in ρ' which is not in ρ. By the property (iii) defining a Dedekind cut, there exists a rational number c in ρ' such that $b < c$. Consequently, we have

$$\rho \, < \, \tau_c \, < \, \rho',$$

where τ_b is the cut representing the rational number b. Done.

Define addition in $\mathbb{R}$ by setting

$$\rho + \rho' \, = \, \{a \in \mathbb{Q} \, | \, \exists \, b \in \rho, c \in \rho' \quad \text{s.t.} \quad a = b + c\}.$$

It is obvious this operation is commutative and associative. Recall that 0 is represented in $\mathbb{R}$ by $\tau_0 = \{x \in \mathbb{Q} | x < 0\}$. Therefore, given any $\rho \in \mathbb{R}$ and $b \in \rho$, $b + x$ being $< b$ for any $x < 0$ implies that $\rho + 0 \leq \rho$. Pick any b in ρ. By the property (iii) of any cut, there is some c in ρ such that $b < c$. Then $c - b$ is a positive rational number, and we can find some rational number $-x$ lying strictly between 0 and $c - b$. Then x is negative and $b < x + c$, which implies that b lies in $\rho + 0$. Thus $\rho + 0$ equals ρ. Thus we have the identities $(F1), (F2)$ and $(F7)$.

For any $\rho \in \mathbb{R}$, one defines its *negative* to be

$$-\rho \, = \, \{-a \in \mathbb{Q} \, | \, a \in \rho^c, \, \exists \, b \in \rho^c \quad \text{s.t.} \quad b < a\},$$

where ρ^c denotes the complement $\mathbb{Q} - \rho$ of ρ in $\mathbb{Q}$. We will leave it to the reader to check that $-\rho$ is indeed a real number, i.e., that it satisfies properties (i)-(iii) of a cut. We claim that $-\rho + \rho \leq 0$. Suppose not. Then there must exist $-a$ in $-\rho$ and x in ρ such that $-a + x \geq 0$, i.e., that $a \leq x$. Then by the property (ii), a must belong to ρ, which is a contradiction because by definition a is in ρ^c. Hence the claim.

Suppose $-\rho + \rho$ is strictly less than 0, there must exist a positive rational number y such that $-y$ is not in $-\rho + \rho$. We claim that there exist x in ρ and $-a$ in $-\rho$ such that

$$x + y + a \, = \, 0.$$

We will treat the case when y lies in ρ, and leave the case when it lies in ρ^c as an exercise. Since ρ is not all of $\mathbb{Q}$ and moreover contains all the rationals to the left of any rational it contains, there must exist an $m \in \mathbb{N}$ such that $my \in \rho$, but $(m+1)y \in \rho^c$. By property (iii), there is some rational number q in ρ with $my < q$. If we put $x = q$ and $-a = (m+1)y + (q - my)$, then $x + y + a$ is zero, as desired. Also, since ρ^c contains $(m+1)y$ which is smaller than $-a$, $-a$ is indeed in $-\rho$, and the claim is proved. Consequently, we also have $(F8)$.

Put

$$\mathbb{R}_{>0} \;=\; \{\rho \,|\, \rho > 0\}$$

and

$$\mathbb{R}_{<0} \;=\; \{\rho \,|\, \rho < 0\}.$$

If ρ in $\mathbb{R}$ contains some positive rational number q, it contains all the rationals less than q and so contains τ_0 properly. Thus ρ is positive, i.e., in $\mathbb{R}_{>0}$. Suppose it contains no positive rational number. If it is non-zero, then it is a proper subset of τ_0, i.e., ρ is in $\mathbb{R}_{<0}$, and moreover, its negative $-rho$ will by definition contain a positive rational number and hence lies in $\mathbb{R}_{>0}$. We get $(Ord1)$ and $\mathbb{R}$ is a disjoint union of $\mathbb{R}_{>0}$, $\{0\}$ and $\mathbb{R}_{<0}$. The verification of the remaining *order axiom* $(Ord2)$ is an easy exercise.

Let us now move to *multiplication*. If ρ, ρ' are positive real numbers, we set

$$(-\rho)(-\rho') \;=\; \rho\rho' \;=\; \mathbb{Q}_{<0} \cup \{q \in \mathbb{Q} \,|\, \exists\, a \in \rho \cap \mathbb{Q}_{>0}, b \in \rho' \cap \mathbb{Q}_{>0} \quad s.t. \quad q = ab\}$$

and

$$(-\rho)\rho' \;=\; \rho(-\rho') \;=\; -(\rho\rho').$$

Evidently, this multiplication is commutative and associative, giving $(F3)$ and $(F4)$. It is left to the reader to verify that $\rho\rho'$ is indeed a real number.

We will now check that $\rho \cdot 1$ is ρ. It suffices to deduce this for positive ρ, in which case the inequality $\rho \cdot 1 \leq \rho$ is clear. Also every non-positive a in ρ lies in $\rho \cdot 1$. So let a be a positive rational number in (the cut defining) ρ. By property (iii), $\exists b \in \rho$ with $a < b$. Put $d = a/b$, which is a rational number < 1 and hence belongs to 1, i.e., to τ_1. Hence $a = bd$ lies in $\rho \cdot 1$, and we have $(F6)$. The proof of the *distributivity axiom* is left to the reader to check.

The *multiplicative inverses* for non-zero real numbers is defined as follows: If $\rho \in \mathbb{R}_{>0}$, set

$$\rho^{-1} \;=\; \mathbb{Q}_{\leq 0} \cup \{b \in \mathbb{Q}_{>0} \,|\, b^{-1} \in \rho^c, \exists\, c \in \rho^c \quad \text{s.t.} \quad c < b^{-1}\}$$

and

$$(-\rho)^{-1} \;=\; -(\rho^{-1}).$$

Again the verification that ρ^{-1} is a real number will be left to the reader. Note that ρ^{-1} is also positive. Let a, b be positive rational numbers lying respectively in ρ and ρ^{-1}. Since by definition b^{-1} is in ρ^c, a must be less than b^{-1}. In other words, ab is < 1, which implies that $ab \in \tau_1$. Hence $\rho\rho^{-1} \leq 1$.

Suppose d is an arbitrary rational number in 1. If $d \leq 0$, it will be in $\rho\rho^{-1}$ and we have nothing to do. So assume that d lies strictly between 0 and 1. Suppose d^{-1} lies in ρ. Then, since $d^{-1} > 1$, there must exist some $n \in \mathbb{N}$ such that $d^{-n} \in \rho$, but d^{-n-1} lies in ρ^c. By property (iii) there is some e in ρ such that $d^{-n} < e$. Put $x = d^{-1}e$, which by virtue of being larger than d^{-n-1}, lies in ρ^c. It is easy to check that x^{-1} lies in ρ^{-1}. Hence $d = ex^{-1}$ lies in $\rho\rho^{-1}$, and we get $\rho\rho^{-1} = 1$, establishing $(F9)$.

Thus we have verified that $\mathbb{R}$ is an ordered field containing $\mathbb{Q}$ and satisfying $(\mathcal{R}1), (\mathcal{R}2)$ and $(\mathcal{R}4)$. It remains to prove that $\mathbb{R}$ is a *complete* ordered field, i.e., verify $(Cord)$, which is the same (for $\mathbb{R}$) as $(\mathcal{R}3)$.

Let X be a bounded set of real numbers. Put

$$\beta = \{a \in \mathbb{Q} \mid \exists \rho \in X \quad \text{s.t.} \quad a \in \rho\}.$$

Properties (i) and (ii) defining a real number are immediate. For property (iii), note that if $a \in \beta$, then by definition $a \in \rho$ for some $\rho \in X$, and since ρ is a real number, there is some b in ρ such that $a < b$. But then b will also belong to β. So β is a real number.

It is clear that β is an upper bound for X. Suppose γ is another upper bound. Then γ will contain every rational number which occurs in any $\rho \in X$. Then γ will necessarily contain every rational number occurring in β. So we get $\beta \leq \gamma$, proving that β is indeed the *least upper bound*.

$\square$

It turns out that $\mathbb{R}$ is essentially the only complete ordered field, the reason for which will not be given here. To be precise we have the following

Theorem *Suppose K is a complete ordered field. Then there is a one-to-one correspondence*

$$f : K \to \mathbb{R},$$

which is compatible with the arithmetical operations on both sides, i.e., $f(0) = 0$, $f(1) = 1$, $f(a + b) = f(a) + f(b)$, and $f(ab) = f(a)f(b)$, for all a, b in K.

1.7 Absolute Value

Given any non-zero real number ρ, we can define its *sign*, denoted $\text{sgn}(\rho)$, to be $+$ (or $+1$) if it is positive, and $-$ (or -1) if it is negative.

Define a function

$$|\,.\,| : \mathbb{R} \to \mathbb{R}$$

by setting

$$|0| = 0$$

and for $\rho \neq 0$,

$$|\rho| = \text{sgn}(\rho)\rho.$$

We will call $|\rho|$ the *absolute value* of ρ.

Proposition *The following hold for all ρ, ρ' in $\mathbb{R}$:*

(*a*)
$$|\rho| \geq 0$$

(*b*)
$$|\rho\rho'| = |\rho|.|\rho'|$$

and

(*c*)
$$|\rho + \rho'| \leq |\rho| + |\rho'|.$$

The inequality (c) is called the *triangle inequality*, signifying that the length of any side of a triangle is less than or equal to the sum of the lengths of the other two sides.

Proof. By definition, if ρ is positive, $|\rho|$, resp. $|-\rho|$, is ρ, resp. $-\rho$, whence (a). The assertion (b) is easy, so we will concentrate on (c). If ρ and ρ' have the same sign, or if one of them is zero, we see that $|\rho + \rho'|$ simply equals $|\rho| + |\rho'|$. So let us assume that they are both non-zero *and* have opposite sign. Interchanging ρ and ρ' if necessary, we may assume that ρ is positive and ρ' is negative. Write $\rho' = -\rho_1$, so that ρ_1 is positive and $|\rho| + |\rho'|$ is $\rho + \rho'$. So we need to show that

$$|\rho - \rho_1| \leq \rho + \rho_1,$$

for positive ρ, ρ_1. The left hand side is $\rho - \rho_1$ if $\rho_1 \leq \rho$, in which case the assertion is clear. So let ρ be less than ρ_1. Then the left hand side is $\rho_1 - \rho$, and again the assertion is clear. □

1.8 Complex Numbers

¿From here on, we will begin to denote real numbers with letters like $x, y, z, \ldots$ near the end of the (Roman) alphabet, instead of the Greek ones. Note that the definitions we gave in section 1.5 are such that for every non-zero real number x, its square $x^2 = x \cdot x$ is always positive. Consequently, $\mathbb{R}$ does not contain the square roots of any negative number. This is a serious problem which rears its head all over the place.

It is a non-trivial fact, however, that any positive number has two square roots in $\mathbb{R}$, one positive and the other negative; the positive one is denoted $\sqrt{x}$. One can show that for any x in $\mathbb{R}$,

$$|x| = \sqrt{x \cdot x}.$$

So if we can somehow have at hand a square root of -1, we can find square roots of any real number.

This motivates us to declare a new entity, denoted i, to satisfy

$$i^2 = -1.$$

One defines the set of *complex numbers* to be

$$\mathbb{C} \;=\; \{x + iy \,|\, x, y \in \mathbb{R}\}$$

and defines the basic arithmetical operations in $\mathbb{C}$ as follows:

$$(x + iy) \pm (x' + iy') \;=\; (x \pm x') + i(y \pm y'),$$

and

$$(x + iy)(x' + iy') \;=\; (xx' - yy') + i(xy' + x'y).$$

There is a natural one-to-one function

$$\mathbb{R} \to \mathbb{C},\; x \to x + i.0,$$

compatible with the arithmetical operations on both sides.

It is an easy exercise to check the field axioms $(F1) - (F8)$ for $\mathbb{C}$. To get the remaining $(F9)$, however, one has to give an argument. To this end one defines the *complex conjugate* of any $z = x + iy$ in $\mathbb{C}$ to be

$$\overline{z} \;=\; x - iy.$$

Clearly,

$$\mathbb{R} \;=\; \{z \in \mathbb{C} \,|\, \overline{z} = z\}.$$

If $z = x + iy$, we have by definition,

$$z\overline{z} \;=\; x^2 + y^2.$$

In particular, $z\overline{z}$ is either 0 or a positive real number. Hence we can find a non-negative square root of $z\overline{z}$ in $\mathbb{R}$. Define the *absolute value*, sometimes called *modulus* or *norm*, by

$$|z| \;=\; \sqrt{z\overline{z}} \;=\; \sqrt{x^2 + y^2}.$$

If $z = x + iy$ is not 0, we will put

$$z^{-1} \;=\; \frac{\overline{z}}{z\overline{z}} \;=\; \frac{x}{x^2 + y^2} - i\frac{y}{x^2 + y^2}.$$

It is a complex number satisfying

$$z(z^{-1}) \;=\; z\frac{\overline{z}}{z\overline{z}} \;=\; 1.$$

Thus $(F9)$ holds in $\mathbb{C}$ as well. Moreover, it is straightforward to check that the Proposition of section 1.6 holds for the absolute value on $\mathbb{C}$. Hence $\mathbb{C}$ is a *field with an absolute value*, just like $\mathbb{R}$.

It is natural to think of complex numbers $z = x + iy$ as being ordered pairs (x, y) of real numbers. So one can try to visualize $\mathbb{C}$ as a plane with two perpendicular coordinate

directions, namely giving the x and y parts. Note in particular that 0 corresponds to the origin $O = (0,0)$, 1 to $(1,0)$ and i with $(0,1)$.

Addition of complex numbers has then a simple geometric interpretation: If $z = x + iy$, $z' = x' + iy'$ are two complex numbers, represented by the points $P = (x,y)$ and $Q = (x',y')$ on the plane, then one can join the origin O to P and Q, and then draw a parallelogram with the *line segments* OP and OQ as a pair of adjacent sides. If R is the fourth vertex of this parallelogram, it corresponds to $z + z'$. This is called the *parallelogram law.*

Complex conjugation corresponds to *reflection* about the x-axis.

The *absolute value* or *modulus* $|z|$ of a complex number $z = x+iy$ is, by the Pythagorean theorem applied to the triangle with vertices $O, P = (x,y)$ and $R = (x,0)$, simply the *length*, often denoted by r, $\sqrt{x^2+y^2}$ of the line OP.

The *angle* θ between the line segments OR and OP is called the *argument* of z. The pair (r,θ) determines the complex number z. Indeed High school trigonometry allows us to show that the coordinates of z are given by

$$x = r\cos\theta \quad \text{and} \quad y = r\sin\theta,$$

where cos (or cosine) and sin (or sine) are the familiar trigonometric functions. Consequently,

$$z = r(\cos\theta + i\sin\theta).$$

Those who know about *exponentials* (to be treated later in the course) will recognize the identity

$$e^{i\theta} = \cos\theta + i\sin\theta.$$

This can also be taken as a definition of $e^{i\theta}$.

Note that $e^{i\theta}$ has absolute value 1 and hence lies on the *unit circle* in the plane given by the equation $|z| = 1$.

2 Sequences and Series

In this chapter we will study two related questions. Given an infinite collection X of numbers, which can be taken to be rational, real or complex, the first question is to know if there is a limit point to which the elements of X congregate (converge). The second question is to know if one can add up all the numbers in X.

Even though we will be mainly thinking of $\mathbb{Q}$, $\mathbb{R}$ or $\mathbb{C}$, many things will go through for numbers in any field K with an absolute value $|\,.\,|$ satisfying $|0| = 0$, $|x| > 0$ if $x \neq 0$, $|xy| = |x| \cdot |y|$, and the *triangle inequality* $|x + y| \leq |x| + |y|$.

A necessary thing to hold for convergence is the *Cauchy condition*, which is also sufficient if we work with $\mathbb{R}$ or $\mathbb{C}$, but not $\mathbb{Q}$. What we mean by this remark is that a Cauchy sequence of rational numbers will have a limit L in $\mathbb{R}$, but L need not be rational. One can enlarge $\mathbb{Q}$ to have this criterion work, i.e., adjoin to $\mathbb{Q}$ limits of Cauchy sequences of rational numbers, in which case one gets yet another construction of $\mathbb{R}$. This aspect is discussed in section 2.3, which one can skip at a first reading.

The last three sections deal with the second question, and we will derive various *tests* for *absolute convergence*. Sometimes, however, a series might converge though not absolutely. A basic example of this phenomenon is given by the *Leibniz series*

$$1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \frac{1}{9} - \frac{1}{11} + \ldots = \frac{\pi}{4}.$$

But the series on the left is *not* absolutely convergent, i.e., the series $1 + \frac{1}{3} + \frac{1}{5} + \ldots$ diverges.

2.1 Convergence of sequences

Let K denote either $\mathbb{Q}$ or $\mathbb{R}$ or $\mathbb{C}$. When we say a *number*, or a *scalar*, we will mean an element of K. But by a *positive number* we will mean a positive *real* number. (One could restrict to positive rational numbers instead, and it will work just as well to use this subset.)

By a *sequence*, or more properly an *infinite sequence*, we will mean a collection of numbers

$$a_1, a_2, a_3, \ldots, a_n, a_{n+1}, \ldots,$$

which is indexed by the set $\mathbb{N}$ of natural numbers. We will often denote it simply as $\{a_n\}$.

A simple example to keep in mind is given by $a_n = \frac{1}{n}$, which appears to decrease towards zero as n gets larger and larger. In this case we would like to have 0 declared as the *limit of the sequence*.

It should be noted that if we are given a collection of numbers indexed by any countably infinite set, we can reindex it by $\mathbb{N}$ and apply all our results below to it. For example, given $M > 0$ and $\{c_n \,|\, n > -M\}$, we can renumber it by putting $a_n = c_{n+M}$ and get a sequence indexed by $\mathbb{N}$.

We will say that a sequence $\{a_n\}$ **converges to a limit** L in K iff for every positive number ϵ there exists an index N such that

$$n > N \implies |L - a_n| < \epsilon. \tag{2.1.1}$$

It is *very important* to note that we have to check the condition above for *every* positive ϵ. Clearly, if (2.2.1) holds for N, it also holds for any $N' > N$.

When $\{a_n\}$ does converge to a limit L, we will write

$$a_n \to L \quad \text{or} \quad \lim_{n\to\infty} a_n = L.$$

When we write $n \to \infty$, we do not necessarily mean that there is a number called ∞, which we have certainly not defined; we mean only that n becomes arbitrarily large. In other words we want to know what happens to a_n when n is larger than any number one can think of.

Let us quickly prove that $\lim\limits_{n\to\infty} \frac{1}{n}$ is 0, as expected. For any positive number ϵ, we can find a rational number N such that $0 < \frac{1}{N} < \epsilon$. Indeed, by the property $\mathcal{R}4$ of $\mathbb{R}$ (see section 1.0), we can find a rational number q with $0 < q < \epsilon$, and since q s a positive rational number, we can find an $N > 0$ such that $\frac{1}{N} < q$. So for every $n > N$, we have

$$|0 - \frac{1}{n}| = \frac{1}{n} < \frac{1}{N} < \epsilon,$$

proving that $\frac{1}{n}$ does converge to 0.

When a sequence does not converge to any number L, we will say that it *diverges*. A simple example is given by the sequence $\{1, -1, 1, -1, 1, -1, \ldots\}$. Given any $\epsilon < 1$, we will be unable to find an N for which (2.1.1) holds for any L. This is so because for any N $\{a_{N+1}, a_{N+2}\} = \{1, -1\}$, and (for any L) at least one of them will be at a distance > 1 from L; to see this use the triangle inequality $|L - a_{N+1}| \le |L - a_{N+2}| + |a_{N+2} - a_{N+1}|$, together with $|a_{N+2} - a_{N+1}| = 2$.

Any sequence which converges (to some limit) will be called a *convergent sequence*, and one that does not converge will be called a *divergent sequence*.

Proposition 2.1.2

(i) If $\{a_n\}$ is a convergent sequence with limit L, then for any scalar c, the sequence $\{ca_n\}$ is convergent with limit cL;

(ii) If $\{a_n\}$, $\{b_n\}$ are convergent sequences with respective limits L_1, L_2, then their sum $\{a_n + b_n\}$ and their product $\{a_nb_n\}$ are convergent with respective limits $L_1 + L_2$ and L_1L_2.

The proof is a simple application of the properties of absolute values. To illustrate, let us prove the assertion about the sum sequence $\{a_n + b_n\}$. Pick any $\epsilon > 0$. Since $a_n \to L_1$ and $b_n \to L_2$, we can find $N > 0$ such that

$$n > N \implies |L_1 - a_n| < \frac{\epsilon}{2}, |L_2 - b_n| < \frac{\epsilon}{2}.$$

Using the triangle inequality we then have, for all $n > N$,

$$|(L_1 + L_2) - (a_n + b_n)| \le |L_1 - a_n| + |L_2 - b_n| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$

Hence $a_n + b_n \to L_1 + L_2$.

Corollary 2.1.3 *Let $\{z_n = x_n + iy_n\}$ be a sequence of complex numbers, with x_n, y_n real for each n. Then $\{z_n\}$ converges iff the real sequences $\{x_n\}$ and $\{y_n\}$ are both convergent.*

Proof. Suppose $\{x_n\}, \{y_n\}$ are both convergent, with respective limits u, v. We claim that $\{z_n\}$ then converges to $w = u+iv$. Indeed, by the Proposition above, $\{iy_n\}$ is convergent with limit iv, and so is $\{x_n + iy_n\}$, with limit w. Conversely, suppose that $\{z_n\}$ converges, say to w. We may write w as $u + iv$, with u, v real. For any complex number $z = x + iy$, $|x|$ and $|y|$ are both bounded by $\leq \sqrt{x^2 + y^2}$, i.e., by $|z|$. Since $w - z_n = (u - x_n) + i(v - y_n)$, we get

$$|u - x_n| \;\leq\; |w - z_n| \quad \text{and} \quad |v - y_n| \;\leq\; |w - z_n|.$$

For any $\epsilon > 0$, pick $N > 0$ such that for all $n > N$, $|w - z_n|$ is $< \epsilon$. Then we also have $|u - x_n| < \epsilon$ and $|v - y_n|$ is $< \epsilon$ for all $n > N$, establishing the convergence of $\{x_n\}$ and $\{y_n\}$ with respective limits u and v.

□

Thanks to this Corollary, the study of sequences of complex numbers can be reduced to that of real numbers.

A sequence $\{a_n\}$ is *bounded* iff there is a positive number M such that $|a_n| \leq M$ for all n.

Lemma 2.1.4 If a sequence $\{a_n\}$ converges, then it must be bounded.

Proof Let L be the limit of $\{a_n\}$. Fix any positive number ϵ. Then by definition, there exists an $N > 0$ such that for all $n > N$, $|L - a_n|$ is $< \epsilon$. By the triangle inequality,

$$|a_n| \;\leq\; |a_n - L| + |L|.$$

So if we set

$$M \;=\; \max\{a_1, a_2, \ldots, a_N, L + \epsilon\},$$

then every $|a_n|$ is bounded by M.

□

The converse of this Lemma is *not* true: Remember that the sequence $\{1, -1, 1, -1, \ldots\}$ has not limit, but it is bounded in absolute value by 1.

For any real number x and consider the *power sequence* $\{x^n\}$. When $x = 1$ (resp. 0), we get the *stationary sequence* $\{1, 1, 1, \ldots\}$ (resp. $\{0, 0, 0, \ldots\}$), which is obviously convergent with limit 1 (resp. 0). When $x = -1$, we get $\{-1, 1, -1, \ldots\}$, which is divergent, as we saw above. Applying the following Lemma we see that this sequence is unbounded when $|x| > 1$, hence divergent, and that it converges to zero when $|x| < 1$.

Lemma 2.1.5 *Let y be a positive real number. Then $\{y^n\}$ is unbounded if $y > 1$, while*

$$\lim_{n\to\infty} y^n \;=\; 0 \quad \text{if} \quad y < 1.$$

Proof of Lemma. Suppose $y > 1$. Write $y = 1 + t$ with $t > 0$. Then by the binomial theorem (which can be proved by induction),

$$y^n = (1+t)^n = \sum_{k=1}^{n} \binom{n}{k} t^k,$$

which is $\geq 1 + nt$. Since $1 + nt$ is unbounded, i.e, larger than any number for a big enough n, y^n is also unbounded.

Now let $y < 1$. Then y^{-1} is > 1 and hence $\{y^{-n}\}$ is unbounded. This implies that, for any $\epsilon > 0$, y^n is $< \epsilon$ for large enough n. Hence the sequence y^n converges to 0.

□

A sequence of real (or rational) numbers a_n is said to be *monotone increasing*, resp. *monotone decreasing*, iff for all n, $a_n \leq a_{n+1}$, resp. $a_n \geq a_{n+1}$; it is said to be simply *monotone* if it is either monotone increasing or monotone decreasing.

Theorem A *If $\{a_n\}$ is a* **bounded, monotone** *sequence of* **real** *numbers, then it converges in $\mathbb{R}$.*

Proof. Suppose $\{a_n\}$ is monotone *increasing*. Since $\mathbb{R}$ is a complete ordered field, every bounded subset has a least upper bound. Since $\{a_n\}$ is bounded, we may look at its supremum M, say. We claim that $a_n \to M$. Suppose not. Then there exists an $\epsilon > 0$ such that for every $N > 0$, there exists an $n > N$ such that $|M - a_n| \geq \epsilon$. Since the sequence is monotone increasing, we then have

$$|a_n| < M - \frac{\epsilon}{2}, \quad \forall n \geq 1.$$

Then $M - \frac{\epsilon}{2}$ is an upper bound for $\{a_n\}$, and so M could not be the *least* upper bound. Contradiction! So M must be the limit of the sequence.

Now suppose $\{a_n\}$ is monotone *decreasing*. Then the sequence $\{-a_n\}$ is monotone increasing, and it is bounded because $\{a_n\}$ is. So $\{-a_n\}$ converges; then so does $\{a_n\}$.

□

Here is an example. Define a sequence $\{a_n\}$ by putting

$$a_n = 1 + \frac{1}{1!} + \frac{1}{2!} + \ldots + \frac{1}{(n-1)!}.$$

It is not hard to see that this sequence is bounded. Try to give a proof. (In fact one can show that it is bounded by 3, but we do not need the best possible bound at this point.) Clearly, $a_{n+1} > a_n$, so the sequence is also monotone increasing. So we may apply the Theorem and conclude that it converges to a limit e, say, in $\mathbb{R}$. But it should be remarked that one can show with more work that e is irrational. So there is a valuable lesson to be learned here. Even though $\{a_n\}$ is a bounded, monotone sequence of rational numbers, there is no limit in $\mathbb{Q}$; one has to go to the enlarged number system $\mathbb{R}$.

A *subsequence* of a sequence $\{a_n\}$ is a subcollection indexed by a countably infinite subset J of $\mathbb{N}$. For example, $\{a_{2n}\}$ and $\{a_{2n+1}\}$ are subsequences.

Theorem B *Suppose $\{a_n\}$ is a bounded sequence of real numbers. Then it has a subsequence which converges.*

Proof. Put

$$J = \{m \in \mathbb{N} \,|\, a_n \leq a_m \,\forall\, n \geq m\}.$$

Suppose J is an infinite set. Write its elements as $m_1, m_2, m_3, \ldots$. By definition, we have

$$a_{m_1} \geq a_{m_2} \geq am_3 \geq \ldots.$$

So the subsequence $\{a_m \,|\, m \in J\}$ (of $\{a_n\}$) is a bounded, monotone decreasing sequence, hence convergent by Theorem A. We are done in this case.

Now suppose J is finite. Pick an n_1 which is larger than any element of J, which is possible. Then we can find an $n_2 > n_1$ such that $a_{n_1} < a_{n_2}$. Again we ca find an n_3 such that $a_{n_2} < a_{n_3}$, and so on. This way we get a subsequence $\{a_{n_1}, a_{n_2}, a_{n_3}, \ldots\}$ which is bounded and monotone increasing, hence convergent by Theorem A.

□

2.2 Cauchy's criterion

The main problem with the definition of convergence of a sequence is that it is hard to verify it. For instance we need to have a candidate for the limit to verify (2.1.2). One wants a better way to check for convergence. As seen above, boundedness is a necessary, but not sufficient, condition, unless the sequence is also monotone (and real). Is there a necessary and sufficent condition? A nineteenth century French mathematician named Cauchy gave an affirmative answer for real sequences. His criterion also works for complex sequences.

A sequence $\{a_n\}$ is said to be a *Cauchy sequence* iff we can find, every positive ϵ, an $N > 0$ such that

$$(2.2.1) \qquad |a_m - a_n| < \epsilon \quad \text{whenever} \quad n, m > N.$$

This is nice because it does not mention any limit.

Lemma 2.2.2 *Every convergent sequence is Cauchy.*

Proof. Suppose $a_n \to L$. Pick any $\epsilon > 0$. Then by definition, we can find an $N > 0$ such that for all $n > N$, we have

$$|L - a_n| < \frac{\epsilon}{2}$$

Then for $n, m > N$, the triangle inequality gives

$$|a_m - a_n| \leq |a_m - L| + |L - a_n| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$

□

Lemma 2.2.3 *A complex sequence $\{z_n = x_n + iy_n\}$, with x_n, y_n real, is Cauchy iff $\{x_n\}$ and $\{y_n\}$ are convergent.*

Proof. Suppose $\{x_n\}, \{y_n\}$ are both Cauchy. Pick any $\epsilon > 0$. Then we can find an $N > 0$ such that for all $m, n > N$,

$$|x_m - x_n| < \frac{\epsilon}{2}, \quad \text{and} \quad |y_m - y_n| < \frac{\epsilon}{2}.$$

Since $z_m - z_n$ is $(x_m - x_n) + i(y_m - y_n)$ and $|i| = 1$, , we get by the triangle inequality,

$$|z_m - z_n| \leq |x_m - x_n| + |y_m - y_n| < \epsilon.$$

Conversely, suppose that $\{z_n\}$ is Cauchy. Recall that for any complex number $z = x+iy$, $|x|$ and $|y|$ are both bounded by $\leq \sqrt{x^2 + y^2}$, i.e., by $|z|$. This immediately implies that $\{x_n\}$ and $\{y_n\}$ are Cauchy.

□

Here comes the beautiful result of Cauchy:

Theorem C *Every Cauchy sequence of real (resp. complex) numbers converges in* $\mathbb{R}$ *(resp.* $\mathbb{C}$*).*

Proof. By Corollary 2.1.3, a sequence $\{z_n = x_n + iy_n\}$ of complex numbers is convergent iff the real sequences $\{x_n\}$ and $\{y_n\}$ are convergent. By Lemma 2.2.3, $\{z_n\}$ is Cauchy iff $\{x_n\}, \{y_n\}$ are. So it suffices to prove Theorem C for *real* Cauchy sequences $\{a_n\}$.

First note that such a sequence must be *bounded*. Indeed, fix an $\epsilon > 0$ and find an $N > 0$ for which (2.2.1) holds. Then for all $n > N$, a_n lies between $a_{N+1} - \epsilon$ and $a_{N+1} + \epsilon$. Then the set $\{a_n \,|\, n > N\}$ is bounded. Then so is the entire sequence because $\{a_1, \ldots, a_N\}$ is finite.

Now since $\{a_n\}$is a bounded, real sequence, we can apply Theorem B and find a subsequence $\{a_{m_1}, a_{m_2}, \ldots\}$ which converges, say to the limit L. We claim that the original sequence $\{a_n\}$ also converges to L. Indeed, let ϵ be an arbitrary positive number. Then we can find an $N > 0$ such that for all $m_k, n > N$,

$$|L - a_{m_k}| < \frac{\epsilon}{2} \quad \text{and} \quad |a_{m_k} - a_n| < \frac{\epsilon}{2}.$$

Applying the triangle inequality we get

$$|L - a_n| < \epsilon \quad \forall n > N,$$

proving the claim.

□

2.3 Construction of Real Numbers revisited

Given any *Cauchy sequence* $\{a_n\}$ **of rational numbers**, we may of course view it as a real sequence and apply Theorem C to deduce that it converges to a limit L, say. But this L need not be rational. This leads us to wonder if *all* the real numbers could be constructed by adjoining limits of Cauchy sequences of rational numbers. This is essentially correct, and

this approach is quite different from that given in chapter 1. Now let us explain how it is done.

Let us begin by pretending that we do not know what real numbers are, but that we know everything about rational numbers.

The first thing to note is that different Cauchy sequences of rational numbers might have the same limit. For example, the sequences $\{0, 0, 0, \ldots\}$ and $\{1, \frac{1}{2}, \frac{1}{3}, \ldots\}$ have 0 as their limit. So we define an **equivalence** between any pair of Cauchy sequences $\{a_n\}$ and $\{b_n\}$ with $a_n, b_n \in \mathbb{Q}$ are as follows:

$$\{a_n\} \sim \{b_n\} \Leftrightarrow \lim_{n\to\infty} (a_n - b_n) = 0. \tag{2.3.1}$$

It will be left as a good exercise for the students to check that $\sim$ is indeed an equivalence relation in the sense of chapter 0.

Denote the equivalence class of any $\{a_n\}$ by $[a_n]$. It is not hard to check that given two classes $[a_n]$ and $[b_n]$, they are either the same or disjoint, i.e., no rational Cauchy sequence $\{c_n\}$ can belong to two distinct classes.

Define a **real number** to be an equivalence class $[a_n]$ as above. Denote by $\mathbb{R}$ the collection of all real numbers in this sense. We then have a natural, mapping

$$\mathbb{Q} \to \mathbb{R}, \ q \to [q, q, q, \ldots],$$

where $[q, q, q, \ldots]$ is the class of the **stationary sequence** $\{q, q, q, \ldots\}$ (which is evidently Cauchy). This mapping is one-to-one because if q, q' are unequal rational numbers, then $\{q, q, q, \ldots\}$ cannot be equivalent to $\{q', q', q', \ldots\}$.

Consequently, we have 0 and 1 in $\mathbb{R}$.

We know that the sum and product of Cauchy sequences are again Cauchy. So we may set

$$[a_n] \pm [b_n] = [a_n \pm b_n], \quad \text{and} \quad [a_n] \cdot [b_n] = [a_n b_n].$$

It is immediate that one has commutativity, associativity and distributivity.

It remains to define a multiplicative inverse of a non-zero element ρ of $\mathbb{R}$. Let $\{a_n\}$ be a Cauchy sequence belonging to ρ. Since ρ is non-zero, all but a finite number of the a_n must be non-zero. Let $\{a_{m_n}\}$ be the subsequence consisting of all the non-zero terms of $\{a_n\}$. Then this subsequence also represents ρ, and we may set

$$\rho^{-1} = [a_{m_n}^{-1}].$$

Check that this gives a well defined multiplicative inverse.

Now it is more or less obvious that $\mathbb{R}$ is a field under these operations, containing $\mathbb{Q}$ as a subfield.

Next comes the notion of *positivity*. Define a real number ρ to be positive iff for some Cauchy sequence $\{a_n\}$ representing ρ, we can find an $N > 0$ such that a_n is positive for all $n > N$. Check that this intuitive notion is the right one.

Armed with this powerful notion, one sees that $\mathbb{R}$ becomes an **ordered field**. By definition, the ordering on $\mathbb{R}$ is compatible with the usual one on $\mathbb{Q}$.

Finally, every Cauchy sequence converges in $\mathbb{R}$. This is a nice consequence of our construction. When ρ is represented by a Cauchy sequence $\{a_n\}$, the sequence is forced to converge to ρ.

Now we will end this section by discussing how one can mathematically define real numbers via their **decimal expansions**, which everyone intuitively understands. Here is what one does.

Define a **real number** to be a pair $(m, \{a_n\})$, where m is an integer and $\{a_n\}$ is a sequence of non-negative integers, with the understanding that a_n is not all 9 after some index N. One often writes

$$(m, \{a_n\}) \;=\; m + \sum_{n=1}^{\infty} a_n 10^{-n}.$$

If we have a rational number of the form $m + \sum\limits_{n=1}^{r} a_n$, we will identify it with $(m, \{a_n\})$ with $a_n = 0$ for all $n > r$.

Define an ordering by setting

$$(m, \{a_n\}) \;<\; (k, \{b_n\}) \;\Leftrightarrow\; m < k \quad \text{or} \quad m = k, \exists N \,\text{s.t.}\, a_N < b_N,\; a_n = b_n \,\forall\, n < N.$$

This leads quickly to the **least upper bound property**, so we can take the suprema of bounded sets.

Given any real number $\rho = (m, \{a_n\})$, we will define, for each $r \geq 1$, its **truncation** at stage r to be

$$\rho_r \;=\; m + \sum_{n=1}^{r} a_n \;\in\; \mathbb{Q}.$$

Evidently,

$$\rho \;=\; \sup\{\rho_r \,|\, r \geq 1\}.$$

So it is natural to set, for real numbers ρ, ρ',

$$\rho \pm \rho' \;=\; \sup\{\rho_r \pm \rho'_r\}$$

and

$$\rho \cdot \rho' \;=\; \sup\{\rho_k \cdot \rho'_k\}.$$

The definition of the multiplicative inverse of a non-zero ρ is left as an exercise.

2.4 Infinite series

Given a sequence $\{a_n\}$, we will often want to know if we can sum the terms of the sequence, denoted

(2.4.1)
$$S \;=\; \sum_{n=1}^{\infty} a_n,$$

and get something meaningful. Such an infinite sum is called an *infinite series*.

A natural way to proceed is as follows. We know how to sum, or any $n \geq 1$, the first n terms of the sequence, namely

$$s_n = \sum_{m=1}^{n} a_n. \tag{2.4.2}$$

This way we get another sequence $\{s_n\}$, called the sequence of *partial sums*. (One calls s_n the nth partial sum.) We will say that the original sequence $\{a_n\}$ is *summable*, or that the infinite series (2.4.1) is convergent, iff the sequence of partial sums converges.

Let us first look at the example of a **geometric series**. Fix any real number x with $|x| < 1$, and put

$$a_n = x^{n-1} \, (\forall \, n \geq 1),$$

and

$$S = \sum_{n=1}^{\infty} x^{n-1} = 1 + x + x^2 + x^3 + \ldots \tag{2.4.3}$$

We claim that

$$s_n = \frac{1 - x^n}{1 - x}.$$

Indeed, we have

$$(1-x)s_n = (1-x)(1+x+x^2+\ldots+x^{n-1}) = (1-x)+(x-x^2)+\ldots+(x^{n-1}-x^n) = 1-x^n,$$

proving the claim.

Now by Lemma 2.1.5 (and the discussion preceding it), we have

$$x^n \to 0 \quad \text{since} \quad |x| < 1.$$

Consequently,

$$S = \lim_{n\to\infty} s_n = \frac{1}{1-x}. \tag{2.4.2}$$

Check that the series is divergent if $|x| \geq 1$.

Lemma 2.4.3 *If $\{a_n\}, \{b_n\}$ are summable sequences, then so is $\{c_1 a_n + c_2 b_n\}$ for any pair of scalars c_1, c_2.*

The proof is straight-forward and will be left as an exercise. The reader should take note, however, that nothing is claimed here about the product of two sequences.

We will say that a series $\sum_{n\geq 1} a_n$ is *absolutely convergent* if the series $\sum_{n\geq 1} |a_n|$ is convergent. In this case one also says that the sequence $\{a_n\}$ is *absolutely summable*. This notion is important and makes us want to understand well the summability of positive sequences,

which will be our focus in the next section. It should be noted however (see section 2.6) that there are summable sequences, which are not absolutely summable.

One of the most important reasons for studying absolutely summable sequences is that it allows one to study the product of two sequences. This is encapsulated in the following Proposition, which we will quote without proof.

Proposition 2.4.4 *Suppose $\{a_n\}$, $\{b_n\}$ are two absolutely summable sequences. Consider the sequence $\{c_n\}$ with*

$$c_n = \prod_{i+j=n} a_i b_j.$$

Then $\{c_n\}$ is absolutely summable and moreover,

$$\sum_{n=1}^{\infty} c_n = \sum_{n=1}^{\infty} a_n \cdot \sum_{n=1}^{\infty} b_n.$$

2.5 Tests for Convergence

We will concentrate mainly on the summability of sequences with *non-negative* terms.

The first thing one notices about such a sequence $\{a_n\}$ with each a_n non-negative, is that

$$s_1 \leq s_2 \leq \ldots \leq s_n \leq \ldots,$$

i.e., $\{s_n\}$ is monotone increasing. In view of Theorem A, we then get the following

Lemma 2.5.1 *A sequence of non-negative numbers is summable iff the associated sequence $\{s_n\}$ of partial sums is bounded.*

The first test for convergence is given by the following

Proposition 2.5.2 (Comparison test) *Let $\{a_n\}$, $\{b_n\}$ be non-negative sequences of real numbers, such that*

(i) $a_n \leq b_n$ for all $n \geq 1$;   and

(ii) The sequence $\{b_n\}$ is summable.

Then $\{a_n\}$ is summable.

Proof. Let s_n, t_n denote the n-th partial sums of $\{a_n\}$, $\{b_n\}$ respectively. Then we have (for all n

$$s_n \leq t_n.$$

Since by hypothesis the sequence $\{b_n\}$ is summable, the sequence $\{t_n\}$ converges. Then the t_n, and hence the s_n, must be bounded. So the sequence $\{a_n\}$ is summable by Lemma 2.5.2. □

Here comes the second test.

Proposition 2.5.4 (Ratio test)) *Let $\{a_n\}$ be a sequence of non-negative terms. Suppose the sequence $\{a_{n+1}/a_n\}$ of ratios converges to a real number r.*

(i) If $r < 1$, then $\{a_n\}$ is summable;

(ii) If $r > 1$, then $\{a_n\}$ is not summable, i.e., the series $\sum_{n\geq 1} a_n$ diverges.

It is important to note that this test says nothing when $r = 1$.

Proof. Suppose r is less than 1. Pick any $\epsilon > 0$ with $r + \epsilon < 1$. Write $t = r + \epsilon$, so that $r < t < 1$. Then as $\{a_{n+1}/a_n\}$ converges to r, we can find an $N > 0$ such that for all $n > N$, $|r - a_{n+1}/a_n| < \epsilon$. This implies that

$$n > N \Rightarrow \frac{a_{n+1}}{a_n} < t.$$

In other words,

$$a_{N+2} < a_{N+1}t,\ a_{N+3} < a_{N+2}t < a_{N+1}t^2, \ldots$$

We can rewrite it as saying that for all $j \geq 1$, we have

$$a_{N+j} < a_{N+1}t^{j-1},$$

so that

$$\sum_{j=1}^{\infty} a_{N+j} < a_{N+1} \sum_{j\geq 1}^{\infty} t^{j-1}.$$

This converges by the comparison test, because the geometric series $\sum_{j\geq 1} t^{j-1}$ occurring on the right is convergent as t is positive and < 1 (with limit $\frac{1}{1-t}$). Consequently, the (partial) series $\sum_{n>N} a_n$, and hence the whole series $\sum_{n\geq 1} a_n$, converges.

When $r > 1$, one sees easily that the sequence of partial sums is not bounded, hence divergent.

$\square$

Given a positive real number x and a positive integer n, there is a unique positive real number, denoted $x^{1/n}$, which is an n-th root of x, i.e., $(x^{1/n})^n = x$. We will also set $0^{1/n} = 0$.

Proposition 2.5.5 (Root test) *Let $\{a_n\}$ be a sequence of non-negative real numbers. Suppose the associated root sequence $\{a_n^{1/n}\}$ is convergent with limit R, Then the following hold:*

(i) If $R < 1$, then the series $\sum_{n\geq 1} a_n$ converges;

(ii) If $R > 1$, then the series $\sum_{n\geq 1} a_n$ diverges.

Just as in the Ratio test, this test says nothing when $R = 1$.

Proof. Suppose $R < 1$. Pick any $\epsilon > 0$ with $R + \epsilon < 1$. Write $t = R + \epsilon$, so that $R < t < 1$. Then as $\{a_n^{1/n}\}$ converges to R, we can find an $N > 0$ such that for all $n > N$, $|R - a_n^{1/n}| < \epsilon$. This implies that

$$n > N \Rightarrow a_n < t^n.$$

Hence

$$\sum_{n>N} a_n < \sum_{n>N} t^n < \sum_{n\geq 1} t^n.$$

The (partial) series on the left now converges by the comparison test, because the geometric series $\sum_{n\geq 1} t^n$ occurring on the right is convergent (as t is positive and < 1). Consequently, the complete series $\sum\limits_{n\geq 1} a_n$ also converges.

When $R > 1$, the sequence of partial sums is not bounded, hence divergent.

□

2.6 Alternating series

One of the most beautiful formulas in Mathematics is the formula of Leibniz for $\pi/4$ given at the beginning of this chapter. It is one of the glorious examples of what are known as *alternating series*, i.e., whose terms alternate in sign, which converge but not absolutely. Some people use the terminology *conditional convergence* for this phenomenon. Leibniz was amazing in that he not only found a slew of such formulas, he also discovered a theorem which gives precise conditions on an alternating series to be convergent. Let us now state and prove it.

Theorem D (*Leibniz*) *Let $\{b_n\}$ be a sequence of* **non-negative**, **monotone decreasing** *sequence such that*

$$\lim_{n\to\infty} b_n = 0.$$

Then the alternating series

$$S = \sum_{n=1}^{\infty} (-1)^{n+1} b_n = b_1 - b_2 + b_3 - b_4 + \ldots$$

is convergent.

Proof. Put $a_n = (-1)^{n+1} b_n$ and

$$s_n = a_1 + a_2 + \ldots + a_n = b_1 - b_2 + \ldots + (-1)^{n+1} b_n.$$

We have to show that the sequence $\{s_n\}$ of partial sums converges. First note that for any $n \geq 1$,

$$s_{2n} = s_{2n-1} - b_{2n} \leq s_{2n-1}, \tag{2.6.1}$$

since the b_j are non-negative. We also have (for all $n \geq 2$)

$$s_{2n} = s_{2n-2} + b_{2n-1} - b_{2n} \leq s_{2n-2}, \tag{2.6.2}$$

and

$$s_{2n-1} = s_{2n-3} - b_{2n-2} + b_{2n-1} \geq s_{2n-3}, \tag{2.6.3}$$

both since the b_j are monotone decreasing.

Now let i, j be natural numbers such that i is even and j odd. Then we may pick an $n > 0$ such that $2n \geq i$ and $2n - 1 \geq j$. Applying (2.6.1), (2.6.2) and (2.6.3), we deduce that

$$s_i \leq s_{2n} \leq s_{2n-1} \leq s_j. \tag{2.6.4}$$

In particular, s_i is bounded from above, for any even index i, by $s_1 = b_1$. Similarly, for any odd index j, s_j is bounded from below by $s_2 = b_1 - b_2$.

The inequalities given by (2.6.2) and (2.6.3) show that the subsequence $\{s_{2n}\}$ is monotone increasing and $\{s_{2n-1}\}$ is monotone decreasing. Since the former is bounded from above, and the latter from below, they both converge, say to respective limits ℓ and L, i.e., we have

$$\lim_{n\to\infty} s_{2n} = \ell \leq L = \lim_{n\to\infty} s_{2n-1}.$$

The assertion $\ell \leq L$ follows from (2.6.4).

Finally, note that by definition,

$$s_{2n} - s_{2n-1} = -b_{2n}.$$

Since the b_n converges to zero as $n \to \infty$, we see that the subsequences $\{s_{2n}\}$ and s_{2n-1} must both converge to the same limit, i.e., $\ell = L$. Consequently,

$$\lim_{n\to\infty} s_n = L.$$

□

Here are some examples of alternating series:

$$S_1 = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \dots \tag{2.6.5}$$

$$S_2 = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \dots \tag{2.6.6}$$

$$S_3 = \pi - \frac{\pi^2}{3!} + \frac{\pi^5}{5!} - \frac{\pi^7}{7!} + \dots \tag{2.6.7}$$

Clearly, $\{\frac{1}{n}\}$ and $\{\frac{1}{2n-1}\}$ are non-negative, decreasing sequences converging to zero. So we may apply Leibniz's theorem above and conclude that S_1 and S_2 converge. Even though we will not evaluate these limits here, let us note that

$$S_1 \;=\; \log 2 \quad \text{and} \quad S_2 \;=\; \frac{\pi}{4}. \tag{2.6.8}$$

Let us now turn to (2.6.7). We need to know that

$$\frac{\pi^n}{n!} \;\leq\; \frac{\pi^{n-1}}{(n-1)!} \tag{2.6.9}$$

and that

$$\frac{\pi^n}{n!} \;\to\; 0 \quad \text{as} \quad n \to \infty. \tag{2.6.10}$$

Since π is known to be < 4, (2.6.9) is satisfied for all $n \geq 4$, and when we study exponential functions later in the course, we will see the truth of (2.6.10). So by Leibniz's theorem, the series starting with $n = 4$ converges. So S_3 itself converges. The amazing things is that one can show (to come later)

$$S_3 \;=\; 0.$$

3 Basics of Integration

In this chapter we will study integration of reasonable functions over nice subsets A of $\mathbb{R}$. We will not consider here the integration of complex-valued functions or the integration over subsets of $\mathbb{C}$; that is in essence a part of *vector calculus* (Ma1c material), since $\mathbb{C}$ identifies with the *plane* $\{(x, y) \,|\, x, y \in \mathbb{R}\}$.

Before beginning our rigorous treatment of integration, which might be a bit abstract for some, here is the intuitive idea behind it. Given any function $f : A \to \mathbb{R}$, $A \subset \mathbb{R}$, we can consider its **graph** defined to be

$$\Gamma(f) = \{(x, y) \,|\, x \in A, y = f(x)\},$$

which is typically a curve in the plane. If f is non-negative on A, the graph will lie above the (horizontal) x-axis, though it might touch the x-axis at places. We would like to define the **integral of** f **over** A to be the **area**, if it makes sense, of the region between the graph of f and the part of the x-axis lying in A. It is helpful to first think of A as being a closed interval $[a, b]$, i.e, the set of points x in $\mathbb{R}$ with $a \leq x \leq b$. (Necessarily, $a < b$.) Then the area we want will be bounded by the graph Γ_f, the x-axis, and the vertical lines $\{y = f(a)\}$ and $\{y = f(b)\}$.

Consider, for example, the squaring function $f(x) = x^2$, with A being the interval $[0, 1]$. (See the figure enclosed as an attachment to this chapter.) How can we define and compute $I(f) = \int_0^1 f(x)dx$? For this we follow a method introduced by the famous nineteenth century German mathematician named Riemann. First note that $f(x)$ is a increasing function on this interval. Roughly speaking (see section 3.2 for a precise description), one chooses numbers $0 < t_1 << t_2 < \ldots < t_n = 1$ and looks at the *upper sum*, which corresponds in the attached figure to the area of the whole hatched area, and the *lower sum*, which corresponds within it, to the area of the darker region. One sees that the desired integral is caught between the upper and lower sums. As n gets larger, the difference Δ_n between the upper and lower sums gets smaller,which means that the hatched areas approximate the integral better. In the limit, as n goes to infinity, this difference becomes zero. In other words, the integral is the limit, as n goes to infinity, of either the upper sum sequence, or the lower sum sequence. For a general function $f(x)$, Δ_n may not converge to zero as n goes to infinity, and in such a case we will say that the integral does not exist. It is important to know that the integral might not exist in a given situation, even though in many practical situations f is continuous, and the integral exists.

3.1 Open, closed and compact sets in $\mathbb{R}$

Given any pair of real numbers a, b with $a < b$, one sets

$$(a, b) = \{x \in \mathbb{R} \,|\, a < x < b\},$$

called the **open interval** with endpoints a, b,

$$[a, b] = \{x \in \mathbb{R} \,|\, a \leq x \leq b\},$$

called the **closed interval** with endpoints a, b,

$$(a, b] = \{x \in \mathbb{R} \,|\, a < x \leq b\},$$

and

$$[a, b) = \{x \in \mathbb{R} \,|\, a \leq x < b\}.$$

The sets $(a, b]$ and $[a, b)$ are called **half-open** (or **half-closed**) intervals.

For $a \in \mathbb{R}$, we will call an interval of the form $(a - r, a + r)$ (for some $r > 0$) an **open interval centered at** a. Some people call it a **basic open set** in $\mathbb{R}$.

Remark. It is important to note that given any pair of open intervals I_1, I_2, we can find an open interval I contained in their intersection $I_1 \cap I_2$.

Given any subset X of $\mathbb{R}$, let us denote by X^c the complement $\mathbb{R} - X$ in $\mathbb{R}$. Clearly, the complement of the empty set $\emptyset$ is all of $\mathbb{R}$.

Let A be a subset of $\mathbb{R}$ and let y be a point in $\mathbb{R}$. Then there are exactly three possibilities for y relative to A:
(3.1.1)

(IP) There exists an open interval I containing y such that $I \subset A$.

(EP) There exists an open interval I centered at y which lies completely in the complement A^c of A.

(BP) Every open interval centered at y meets both A and A^c.

In case (IP), y is called an **interior point of** A. In case (EP), y is called an **exterior point of** A. In case (BP), y is called a **boundary point of** A. Note that in case (IP) $y \in A$, in case (EP) $y \notin A$, and in case (BP) y may or may not belong to A.)

Definition 3.1.2 *A set A in $\mathbb{R}$ is open if and only if every point of A is an interior point.*

Explicitly, this says: "Given any $z \in A$, we can find an open interval I containing z such that $I \subset A$."

Definition 3.1.3 A subset A of $\mathbb{R}$ is closed if its complement is open.

Lemma 3.1.4 *$A \subset \mathbb{R}$ is closed iff it contains all of its boundary points.*

Proof. Let y be a boundary point of A. Suppose y is not in A. Then it belongs to A^c, which is open. So, by the definition of an open set, we can find an open interval I containing y with $I \subset A^c$. Such a I does not meet A, contradicting the condition (BP). So A must contain y.

Conversely, suppose A contains all of its boundary points, and consider any z in A^c. Then z has to be an interior point or a boundary point of A^c. But the latter possibility does not arise as then z would also be a boundary point of A and hence belong to A (by hypothesis). So z is an interior point of A^c. Consequently, A^c is open, as was to be shown.

□

Examples:
(3.1.5)
(1) The empty set ϕ and $\mathbb{R}$ are both open and closed.

Since they are complements of each other, it suffices to check that they are both open, which is clear from the definition.

(2) Let $\{W_\alpha\}$ be a (possibly infinite and perhaps uncountable) collection of open sets in $\mathbb{R}$. Then their union $W = \cup_\alpha W_\alpha$ is also open.

Indeed, let $y \in W$. Then $y \in W_\alpha$ for some index α, and since W_α is open, there is an open set $V \subset W_\alpha$ containing y. Then we are done as $y \in V \subset W_\alpha \subset W$.

(3) Let $\{W_1, W_2, \cdots, W_n\}$ be a **finite** collection of open sets. Then their intersection $W = \bigcap_{j=1}^{n} W_j$ is open.

Proof. Let $y \in W$. Then $y \in W_j$, $\forall j$. Since each W_j is open, we can find an open interval I_j such that $y \in I_j \subset W_j$. Then, by the remark above, we can find an open interval I contained in the intersection of the I_j such that $y \in I$. Done.

Warning: The intersection of an infinite collection of open sets need not be open, as shown by the following (counter)example. Put, for each $k \geq 1$, $W_k = \left(-\frac{1}{k}, \frac{1}{k}\right)$. Then $\cap_k W_k = \{0\}$, which is not open.

(4) Any **finite** set of points $A = \{P_1, \ldots, P_r\}$ is closed.

Proof. For each j, let U_j denote the complement of P_j (in $\mathbb{R}$). Given any z in U_j, we can easily find an open interval V_j containing z which avoids P_j. So U_j is open, for each j. The complement of A is simply $\cap_{j=1}^{r} U_j$, which is then open by (4).

More generally, one can show, by essentially the same argument, that a *finite* union of closed sets is again closed.

It is important to remember that there are many sets A in $\mathbb{R}$ which are neither open nor closed. For example, look at the half-closed, half-open interval $[0, 1)$ in $\mathbb{R}$.

It is customary to extend slightly our notion of an open interval, which some would call a **finite open interval**, and define **infinite open intervals**, for all $a \in \mathbb{R}$, by

$$(a, \infty) = \{x \in \mathbb{R} \mid a < x\}$$

and

$$(-\infty, a) = \{x \in \mathbb{R} \mid x < a\}.$$

At times we will also write $(-\infty, \infty)$ for $\mathbb{R}$.

We are not (at all) claiming here that there are real numbers called ∞ and $-\infty$. They are introduced mainly for notational convenience. It is seen easily that these infinite open intervals are indeed open – use example (2) above.

One also defines **infinite half-open intervals** (for all $a \in \mathbb{R}$) by

$$[a, \infty) = \{x \in \mathbb{R} \mid a \leq x\},$$

and

$$(-\infty, a] = \{x \in \mathbb{R} \,|\, x \le a\}.$$

Since there are no such numbers as ∞ and $-\infty$ in the real number system, it does not make sense to define closed infinite intervals.

We will call a subset Y of $\mathbb{R}$ a **bounded set** iff we can enclose it in a closed interval, i.w., $Y \subset [a, b]$ for some real numbers a, b with $a < b$.

This clearly agrees with the notion of boundedness of sequences encountered in the previous chapter.

Definition 3.1.6 *A subset C of $\mathbb{R}$ which is both closed and bounded is called a* **compact set**.

This definition implies, in particular, that **any closed interval $[a, b]$ is compact**. It will turn out later on that compact sets get sent to compact sets under continuous functions.

3.2 Integrals of bounded functions

We will first discuss the question of integrability of bounded functions on closed intervals, and then move on to integration of continuous functions, and then to integration over more general sets in $\mathbb{R}$. The main tool will be to approximate the integral from above by the **upper sum** and from below by the **lower sum**, relative to various partitions. This method was introduced by the famous nineteenth century German mathematician Riemann, and it is customary to call these sums **Riemann sums**.

To begin, let us define the **length** of a closed interval $[a, b]$ to be

$$\ell([a, b]) = b - a.$$

Note that the *interior* of $[a, b]$ is simply the open interval (a, b).

By a **bounded function** on a set A, we will mean a function

$$f : A \to \mathbb{R},$$

such that the *image* $f(A)$ of f is bounded.

Definition 3.2.1 *A* **partition** *of a closed interval $[a, b]$ is a finite collection P of closed subintervals $J_1, J_2, \ldots, J_r \subseteq [a, b]$ such that*

(i) $[a, b] = \cup_{j=1}^{r} J_j$, and

(ii) the interiors of J_i and J_j have no intersection if $i \neq j$.

Giving such a partition of $[a, b]$ is evidently the same as giving $r+1$ points $t_0, t_1, \ldots, t_{r-1}, t_r$ in $\mathbb{R}$ such that

(3.2.2) $$a = t_0 < t_1 << \ldots < t_{r-1} < t_r = b.$$

The corresponding intervals J_j are obtained by setting

$$J_j \;=\; [t_{j-1}, t_j] \quad \forall\, j \geq 1.$$

Definition 3.2.3 *A* **refinement** *of a partition $P : a = t_o < t_1 < \ldots < t_n = b$ of $[a,b]$ is another partition $P' : a = t'_0 < t'_1 < \ldots < t'_m = b$ with $\{t_i | 1 \leq i \leq n\}$ contained in the set $\{t'_j | 1 \leq j \leq m\}$.*

It is clear from the definition that given any two partitions P, P' of $[a,b]$, we can find a third partition P'' which is simultaneously a refinement of P and of P'. Indeed, if P is given by the set $\{t_i \,|\, 0 \leq i \leq r\}$ and P' by $\{u_j \,|\, 0 \leq j \leq s\}$, P'' can be taken to be defined by the union $\{t_i, u_j \,|\, 0 \leq ir, 0 \leq j \leq s\}$. Such a P'' is called a **common refinement** of P, P'.

Now let f be a bounded function on $[a,b]$, and let $P = \{a = t_0 < t_1 < \ldots < t_n = b\}$ a partition of $[a,b]$. Then f is certainly bounded on each $[T_{j-1}, t_j]$. Remember that it was proved in Chapter 1 that every bounded subset of $\mathbb{R}$ admits a **sup**, the **least upper bound**, and an **inf**, the **greatest lower bound**.

Definition 3.2.4 *The* **upper***, resp.* **lower***,* **sum** *of f over $[a,b]$ relative to the partition $P = \{a = t_0 < t_1 < \ldots < t_n = b\}$ is given by*

$$U(f,P) \;=\; \sum_{j=1}^{r} (t_j - t_{j-1}) \sup(f([T_{j-1}, t_j]))$$

resp.

$$L(f,P) \;=\; \sum_{j=1}^{r} (t_j - t_{j-1}) \inf(f([T_{j-1}, t_j])).$$

Of course we have

$$L(f,P) \leq U(f,P), \qquad \text{for all } P.$$

More importantly, it is clear from the definition that if $P' = \{J'_k\}_{k=1}^m$ is a refinement of P, then

$$L(f,P) \leq L(f,P') \quad \text{and} \quad U(f,P') \leq U(f,P).$$

Put

$$\mathcal{L}(f) = \{L(f,P) \mid P \text{ partition of } [a,b]\} \subseteq \mathbb{R} \tag{3.2.5}$$

and

$$\mathcal{U}(f) = \{U(f,P) \mid P \text{ partition of } [a,b]\} \subseteq \mathbb{R}.$$

Lemma 3.2.6 *$\mathcal{L}(f)$ admits a* **sup***, denoted $\underline{I}(f)$, and $\mathcal{U}(f)$ admits an* **inf***, denoted $\overline{I}(f)$.*

Proof. Thanks to the discussion in Chapter 1, all we have to do is show that $\mathcal{L}(f)$ (resp. $\mathcal{U}(f)$) is bounded from above (resp. below). So we will be done if we show that given any two partitions P, P' of $[a,b]$, we have $L(f,P) \leq U(f,P')$ as then $\mathcal{L}(f)$ will have $U(f,P')$ as an upper bound and $\mathcal{U}(f)$ will have $L(f,P)$ as a lower bound. Choose a third partition P'' which

refines P and P' simultaneously. Then we have $L(f,P) \leq L(f,P'') \leq U(f,P'') \leq U(f,P')$. Done.

$\square$

We always have

$$\underline{I}(f) \leq \overline{I}(f).$$

It is customary to call $\underline{I}(f)$ the **lower integral**, and $\overline{I}(f)$ the **upper integral**, of f over $[a,b]$.

Definition 3.2.7 *A bounded function f on $[a,b]$ is* **integrable** *iff $\underline{I}(f) = \overline{I}(f)$. When such an equality holds, we will simply write $I(f)$ (or $I_{[a,b]}(f)$ if the dependence on $[a,b]$ needs to be stressed) for $\underline{I}(f)$ $(= \overline{I}(f))$, and call it the* **integral of f over** $[a,b]$.

Quite often we will write

$$I(f) = \int_a^b f \quad \text{or} \quad \int_a^b f(x)\,dx.$$

In practice one is loathe to consider *all* partitions P of $[a,b]$. The following lemma tells us something useful in this regard.

Lemma 3.2.8 *Let f be a bounded function on $[a,b]$. Suppose $\{P_n\}$ is an infinite sequence of partitions, with each P_n being a refinement of P_{n-1}, such that the corresponding sequences $\{U(f,P_n\}$ and $\{L(f,P_n)\}$ both converge to a common limit λ in $\mathbb{R}$. Then f is integrable with $\lambda = \int_a^b f(x)dx$.*

The obvious question now is to ask if there are integrable functions. One such example is given by the **constant function** $f(x) = c$, for all $x \in [a,b]$. Then for any partition $P = \{A = t_0 < t_1 < \ldots < t_n = b\}$, we have

$$L(f,P) = U(f,P) = c\sum_{j=1}^{r}(t_j - t_{j-1}) = c(b-a).$$

So $\underline{I}(f) = \overline{I}(f)$ and

$$\int_a^b f = c(b-a).$$

This can be jazzed up as follows.

Definition *A* **step function on** $[a,b]$ *is a function f on $[a,b]$ which is constant on each of the closed subintervals $[t_{j-1}, t_j]$ of* **some** *partition P.*

Lemma 3.2.9 *Every step function f on $[a,b]$ is integrable.*

Proof. By definition, there exists a partition $P = \{a = t_0 < t_1 < \ldots < t_n = b\}$ of $[a, b]$ and scalars $\{c_j\}$ such that $f(x) = c_j$, if $x \in [T_{j-1}, t_j]$. Then, arguing as above, it is clear that for **any** refinement P' of P, we have

$$L(f, P') = U(f, P') = \sum_{j=1}^{n} c_j \, (t_j - t_{j-1}).$$

Hence, $\underline{I}(f) = \overline{I}(f)$.

□

Note that for such a step function f defined by $(P, \{c_j\})$, we have an explicit formula for the integral, namely

$$\int_a^b f(x)dx \;=\; \sum_{j=1}^{n} c_j \, (t_j - t_{j-1}).$$

3.3 Integrability of monotone functions

Let $f : A \to \mathbb{R}$ be a function, with A a subset of $\mathbb{R}$. We will call f **monotone increasing**, resp. **monotone decreasing**, iff we have

$$x_1, x_2 \in A,\, x_1 < x_2 \implies f(x_1) \leq f(x_2) \quad (\text{resp.} \quad f(x_1) \geq f(x_2)).$$

By a **monotone** function, we will mean a function which is either monotone increasing or monotone decreasing.

Lemma 3.3.1 *Let f be a monotone function on a closed interval $[a, b]$. Then f is bounded on $[a, b]$.*

Proof. Suppose f is monotone increasing. Then by definition, $f(a) \leq f(x) \leq f(b)$ for all x in $[a, b]$. Hence f is bounded by $f(a)$ (from below) and $f(b)$ (from above). If f is monotone decreasing, the proof is similar and will be left as an exercise.

□

Theorem 3.3.2 *Let f be a monotone function on $[a, b]$. Then f is integrable.*

Proof. Suppose f is monotone increasing on $[a, b]$. For $n \geq 1$, let $P_n = \{a = t_0 < t_1 < \ldots < t_n = b\}$ be the partition with

$$t_{j+1} - t_j \;=\; \frac{b-a}{n} \quad \forall\, j \leq n-1.$$

Since f is increasing, $f(t_j)$ is, for each j, the *inf* of the values of f on the subinterval $[t_j, t_{j+1}]$, and $f(t_{j+1})$ is the *sup*. Hence

$$L(f, P_n) \;=\; \frac{b-a}{n} \, (f(t_0) + f(t_1) + \ldots + f(t_{n-1}))$$

and

$$U(f, P_n) \; = \; \frac{b-a}{n}\left(f(t_1) + f(t_2) + \ldots + f(t_n)\right).$$

It follows that

$$U(f, P_n) - L(f, P_n) \; = \; \frac{b-a}{n}\left(f(b) - f(a)\right).$$

As n goes to infinity, this difference goes to zero. Hence f is integrable by Lemma 3.2.8.

$\square$

3.4 Computation of $\int_a^b x^s dx$

The function $f(x) = x^m$ is a monotone increasing function for any $m > 0$. So by Theorem 3.3.2, it is integrable on any closed interval $[a, b]$. There is a well known formula for the value of the integral, which can be derived in a myriad of ways. We will give two proofs in this chapter, and here is the first one – due to Riemann, which uses partitions $P = a = t_0 < t_1 < \ldots < t_n = b$ where the subintervals $[t_i, t_{i+1}]$ are not of equal length, but where the ratios t_{i+1}/t_i are kept constant.

Riemann's method, unlike the one we will describe later in section 3.7, is very general, and works also for x^s for *any* real exponent $s \neq -1$, as long as the limits a, b are positive. We have encountered x^s before for rational s, and for those who know about logarithms and exponentials, x^s is defined for any real s and positive x as $e^{s \log x}$; here $\log x$ denotes the *natural logarithm* of x, which some denote by $ln(x)$. It is problematic to define x^s for general s and negative x, except when s is an integer, because numbers like $(-1)^{1/2}$ are not in $\mathbb{R}$.

Proposition 3.4.1 *Let s, a, b be real numbers with $s \neq -1$ and $a < b$. If s is not an integer, assume that a is positive. Then*

$$\int_a^b x^s dx \; = \; \frac{b^{s+1} - a^{s+1}}{s+1}.$$

We will prove this result below only for positive integer values of s, but we will make a remark after the proof about what one needs for the extension to general s.

When $s = -1$, one cannot divide by $s + 1$ and the Proposition cannot hold as stated. It may be useful to note that for any $b > 1$,

$$\int_1^b \frac{1}{x} dx \; = \; \log b.$$

One can take this to be the definition of $\log b$.

Proof of Proposition for positive integral exponents.

We will take s to be a positive integer m. (The assertion is obvious for $m = 0$.) We will also assume, for simplicity of exposition, that $a > 0$, even though the asserted formula holds equally well when $a \le 0$. Write

$$f(x) \;=\; x^m.$$

Put

$$u \;=\; b/a \;>\; 1, \tag{3.4.2}$$

and define, for each $n \ge 1$, a partition

$$P_n :\; a = t_0 < t_1 < t_2 < \ldots < t_{n-1} < t_n = b \tag{3.4.3}$$

such that for each $j \le n$,

$$t_j \;=\; au^{j/n}.$$

Then we have, for all $j < n$,

$$\frac{t_{j+1}}{t_j} \;=\; u^{(j+1)/n - j/n} \;=\; u^{1/n}, \tag{3.4.4}$$

which is independent of j.

The lower sum is given, for each n, by

$$L(f, P_n) \;=\; \sum_{j=0}^{n-1} (t_{j+1} - t_j) t_j^m \;=\; (u^{1/n} - 1) \sum_{j=0}^{n-1} t_j^{m+1}, \tag{3.4.5}$$

where we have used (3.4.4). By the definition of t_j and the fact that $\sum\limits_{j=0}^{n-1} z^j \;=\; \frac{1-z^n}{1-z}$, the expression on the right becomes

$$a^{m+1}(u^{1/n} - 1) \sum_{j=0}^{n-1} u^{j(m+1)/n} \;=\; -a^{m+1} \frac{(1 - u^{1/n})(1 - u^{m+1})}{1 - u^{(m+1)/n}}.$$

Since $-a^{m+1}(1 - u^{m+1})$ equals $b^{m+1} - a^{m+1}$, we get

$$L(f, P_n) \;=\; (b^{m+1} - a^{m+1}) \frac{1 - u^{1/n}}{1 - u^{(m+1)/n}} \;=\; (b^{m+1} - a^{m+1}) \frac{1}{1 + u^{1/n} + u^{2/n} + \ldots + u^{m/n}}. \tag{3.4.6}$$

We know that as $n \to \infty$, $u^{j/n}$ goes to 1 for any fixed j. This shows that

$$\lim_{n \to \infty} \frac{1}{1 + u^{1/n} + u^{2/n} + \ldots + u^{m/n}} \;=\; \frac{1}{m+1}. \tag{3.4.7}$$

Consequently, by (3.4.6),

$$\lim_{n \to \infty} L(f, P_n) \;=\; \frac{b^{m+1} - a^{m+1}}{m+1}. \tag{3.4.8}$$

On the other hand, since $t_{j+1}^m = u^{m/n} t_j^m$, the corresponding upper sum is

$$U(f, P_n) \;=\; \sum_{j=0}^{n-1} (t_{j+1} - t_j)(u^{m/n} t_j^m) \;=\; u^{m/n} L(f, P_n). \tag{3.4.9}$$

And since $u^{m/n} \to 1$ as $n \to \infty$, (3.4.8) and (3.4.9) imply that we have

$$\lim_{n\to\infty} U(f, P_n) \;=\; \lim_{n\to\infty} L(f, P_n) \;=\; \frac{b^{m+1} - a^{m+1}}{m+1}.$$

It follows then (see Lemma 3.2.8) that the (definite) integral of x^m over $[a, b]$ equals this common limit. Incidentally, this computation shows explicitly that $f(x) = x^m$ is integrable and we don't really need to refer to Theorem 3.3.2.

□

Now suppose we want to prove the full force of the Proposition, i.e., treat the case of an arbitrary real exponent $s \neq -1$, by this method. Proceeding as above, we will get

$$L(f, P_n) \;=\; (b^{s+1} - a^{s+1})\phi_s(n), \tag{3.4.6}$$

where

$$\phi_s(n) \;=\; \frac{1 - u^{1/n}}{1 - u^{(s+1)/n}},$$

and

$$U(f, P_n) \;=\; u^{s/n} L(f, P_n).$$

As before, $u^{s/n}$ goes to 1 as $n \to \infty$. So the whole argument will go through if we can establish the following limit:

$$\lim_{n\to\infty} \phi_s(n) \;=\; \frac{1}{s+1}.$$

This can be done, but we will not do it here. In any case, the students should feel free to use the Proposition for all $s \neq -1$.

3.5 Example of a non-integrable, bounded function

Define a function

$$f : [0, 1] \to \mathbb{R} \tag{3.5.1}$$

by the following recipe. If x is irrational, set $f(x) = 0$, and if x is rational, put $f(x) = 1$.

This certainly defines a bounded function on [0.1], and one is led to wonder about the integrability of f.

Proposition 3.5.2 *This f is not integrable.*

Proof. Let P be any partition of $[0, 1]$ given by $0 = t_0 < t_1 < \ldots < t_n = 1$. By the property $\mathcal{R}4$ of $\mathbb{R}$ (see section 1.1), we know that there is a rational number between any

two real numbers. So in every subinterval $[t_j, t_{j+1}]$ there will be some rational number q_j (in fact infinitely many), with $f(q_j) = 1$ by definition. Consequently,

$$U(f, P) = \sum_{j=0}^{n-1} 1 \cdot (t_{j+1} - t_j) = (t_0 - t_1) + (t_1 - t_2) + \ldots + (t_{n-1} - t_n) = 1,$$

because $t_0 = 0$ and $t_n = 1$.

On the other hand, every interval $[c, d]$ in $\mathbb{R}$ must contain an irrational number y. Let us give a proof. If c or d is irrational, then we may take y to be that number, so we can assume that c and d are rational. Then the number $y = c + (d - c)\sqrt{2}/2$ is irrational and lies in $[c, d]$. Consequently, for every j, there is an irrational y_j in $[t_j, t_{j+1}]$, which implies that

$$L(f, P) = 0.$$

Hence

$$U(f, P) - L(f, P) = 1,$$

and this is independent of the partition P. So f is not integrable.

□

It should be noted, however, that there are non-zero integrable, bounded functions f on $[0, 1]$ which are supported on $\mathbb{Q}$, i.e., $f(x)$ is zero for irrational numbers x. But they are not constant on the rational numbers.

Here is a very useful simple fact about rational numbers in the interval $[0, 1]$. For every positive integer n, there are at most n rational numbers whose reduced fraction expressions have n as the denominator. This is also the fact that allows one to establish a one-to-one correspondence between $\mathbb{Q}$ and $\mathbb{N}$, showing that $\mathbb{Q}$ is countable.

3.6 Properties of integrals

The integral $\int_a^b f(x)dx$ is often called a **definite integral** because it has a definite value, assuming that f is integrable. One calls f the **integrand**, a the **lower limit** and b the **upper limit**. It is customary to use the convention

$$\int_b^a f(x)dx = -\int_a^b f(x)dx. \tag{3.6.0}$$

The definite integral has many nice properties as f or $[a, b]$ varies, making our life very pleasant, which we want to discuss in this section.

Proposition 3.6.1 (Linearity in the integrand) *If f, g are integrable over $[a, b]$, so is any linear combination $\alpha f + \beta g$, with $\alpha, \beta \in \mathbb{R}$, and moreover,*

$$\int_a^b \{\alpha f(x) + \beta g(x)\}dx = \alpha \int_a^b f(x) + \beta \int_a^b g(x)dx.$$

Proof. If P is any partition, it is immediate from the definition that

$$L(\alpha f + \beta g, P) = \alpha L(f, P) + \beta L(g, P)$$

and

$$U(\alpha f + \beta g, P) = \alpha U(f, P) + \beta U(g, P).$$

It follows then that

$$\underline{I}(\alpha f + \beta g) = \alpha \underline{I}(f) + \beta \underline{I}(g)$$

and

$$\overline{I}(\alpha f + \beta g) = \alpha \overline{I}(f) + \beta \overline{I}(g).$$

Since f, g are integrable, $\underline{I}(f) = \overline{I}(f)$ and $\underline{I}(g) = \overline{I}(g)$. So the lower and upper integrals of $\alpha f + \beta g$ coincide, proving the assertion.

$\square$

Proposition 3.6.2 (Additivity in the limits) Let a, b, c are real numbers with $a < b < c$, and let f be integrable on $[a, c]$. Then f is integrable on $[a, b]$ and $[b, c]$, and moreover,

$$\int_a^c f(x)dx = \int_a^b f(x)dx + \int_b^c f(x)dx.$$

Proof. Given partitions P, P' of $[a, b], [c, d]$ respectively, $P \cup P'$ defines a partition of $[a, c]$. And if P'' is a partition of $[a, c]$, then we can refine it by adding b to get a partition of the type $P \cup P'$, with P (resp. P') being a partition of $[a, b]$ (resp. $[b, c]$). It follows easily that the lower (resp. upper) integral of f over $[a, c]$ is the sum of the lower (resp. upper) integrals of f over $[a, b]$ and $[b, c]$. The assertion follows.

$\square$

If c is any real number , the function $x \to x + c$ is called **translation** by c. The following Proposition describes the **translation invariance** of the definite integral.

Proposition 3.6.3 *Suppose f is integrable on $[a, b]$ and $c \in \mathbb{R}$. Then the c-***translate** *of f, given by $x \to f(x + c)$, is integrable on $[a - c, b - c]$, and*

$$\int_{a-c}^{b-c} f(x + c)dx = \int_a^b f(x)dx.$$

For any $c \in \mathbb{R}$, the function $x \to cx$ is called the **homothety** (or **stretching**) by c. Some also use the terms **expansion** and **contraction** when $c > 1$ and $0 < c < 1$. The following Proposition describes the **behavior under homothety**.

Proposition 3.6.4 *Suppose f integrable on $[ac, bc]$. Then the function $x \to f(cx)$ is integrable on $[a, b]$ and*

$$\int_{ac}^{bc} f(x)dx = c \int_a^b f(cx)dx.$$

3.7 The integral of x^m revisited, and polynomials

In Proposition 3.4.1 we established the following identity for any $m \geq 0$ and any $[a, b]$:

$$(*) \qquad \int_a^b x^m dx \;=\; \frac{b^{m+1} - a^{m+1}}{m+1}.$$

In view of the linearity property of the integral (see Proposition 3.6.1), we see that if we have any **polynomial**

$$f(x) \;=\; c_0 + c_1 x + \ldots + c_n x^n,$$

then f is integrable over any $[a, b]$. Furthermore, we have the **explicit formula**

$$(**) \qquad \int_a^b f(x)dx \;=\; c_0 x + c_1 \frac{x^2}{2} + \ldots + c_n \frac{x^{n+1}}{n+1}.$$

We will now give an **alternate proof** of $(*)$ for $0 < a < b$.

We first need a Lemma, which is of independent interest. We will call a function $f(x)$ **even**, resp. **odd**, iff $f(-x) = f(x)$, resp. $f(-x) = -f(x)$, for all x. Note that x^j is even if j is even and it is odd if j is odd. Also, $\cos x$ is even while $\sin x$ is odd.

Lemma 3.7.1 *Let $f(x)$ be an integrable function of $[-a, a]$, for some $a > 0$. Then*

$$f \text{ even} \quad \Longrightarrow \int_{-a}^{a} f(x)dx \;=\; 2\int_0^a f(x)dx,$$

$$f \text{ odd} \quad \Longrightarrow \int_{-a}^{a} f(x)dx \;=\; 0.$$

Proof. By Proposition 3.6.4 (and convention (3.6.0)),

$$\int_{-a}^{0} f(x)dx \;=\; \int_0^a f(-x)dx,$$

which equals

$$(-1)^r \int_0^a f(x)dx,$$

with r being 1, resp. -1, when f is even, resp. odd. The lemma now follows by Proposition 3.6.2, which implies that

$$\int_{-a}^{a} f(x)dx \;=\; \int_{-a}^{0} f(x)dx + \int_0^a f(x)dx.$$

$\square$

Now let us begin the alternate proof of (∗). Put

$$I_m = \int_0^1 x^m dx.$$

The key is to show that

$$I_m = \frac{1}{m+1}. \tag{3.7.2}$$

Indeed, we can use Proposition 3.6.4 to deduce that

$$\int_0^b x^m dx = b \int_0^1 (bx)^m dx = b^{m+1} I_m, \tag{3.7.3}$$

and combining this with Proposition 3.6.2, we get

$$\int_a^b x^m dx = \int_0^b x^m dx - \int_0^a x^m dx = (b^{m+1} - a^{m+1}) I_m, \tag{3.7.4}$$

as desired. So let us now prove (3.7.2).

By the translation invariance (Proposition 3.6.3) of definite integrals, we get

$$\int_{-b}^b (x+b)^m dx = \int_0^{2b} x^m dx = 2^{m+1} b^{m+1} I_m. \tag{3.7.5}$$

By the binomial theorem,

$$(x+b)^m = \sum_{j=0}^m \binom{m}{j} x^j b^{m-j}.$$

and so by (3.7.5),

$$2^{m+1} b^{m+1} I_m = \sum_{j=0}^m \binom{m}{j} b^{m-j} \int_{-b}^b x^j dx. \tag{3.7.6}$$

By Lemma 3.7.1, $\int_{-b}^b x^j dx$ equals 0 if j is odd, and twice $\int_0^b x^j dx = 2b^{j+1} I_j$ (by 3.7.3) when j is even.

Since $(*)$ is true for $m = 0$, we can take $m > 0$ and assume, by induction, that $(*)$ holds for all $j < m$. Then we get from the above, the following:

$$2^m I_m \;=\; \sum_{j=0,\, j\,\text{even}}^{m} \binom{m}{j} I_j \;=\; \epsilon_m I_m + \sum_{j=0,\, j\,\text{even}}^{m-1} \binom{m}{j} \frac{1}{j+1}, \tag{3.7.7}$$

where ϵ_m is 1 if m is even and 0 if m is odd. On the other hand,

$$\binom{m}{j}\frac{1}{j+1} \;=\; \frac{m!}{(j+1)!(m-j)!} \;==\; \binom{m+1}{j+1}\frac{1}{m+1},$$

so that

$$\sum_{j=0,\, j\,\text{even}}^{m-1} \binom{m}{j}\frac{1}{j+1} \;=\; \frac{1}{m+1} \sum_{k=0,\, k\,\text{odd}}^{m} \binom{m+1}{k}. \tag{3.7.8}$$

Next we note that for any integer $r \geq 1$,

$$\sum_{k=0,\, k\,\text{odd}}^{r} \binom{r}{k} \;=\; \frac{1}{2}(1-(-1))^r \;=\; 2^{r-1}.$$

Consequently,

$$\sum_{k=0,\, k\,\text{odd}}^{m} \binom{m+1}{k} \;=\; 2^m - \epsilon_m. \tag{3.7.9}$$

Combining (3.7.7.), (3.7.8) and (3.7.9), we get

$$2^m I_m \;=\; \epsilon_m I_m + \frac{2^m}{m+1} - \epsilon_m \frac{1}{m+1}.$$

When m is odd, $\epsilon_m = 0$ and hence

$$2^m I_m \;=\; \frac{2^m}{m+1}.$$

When m is even, $\epsilon_m = 1$ and we get

$$2^m I_m \;=\; I_m + \frac{2^m}{m+1} - \frac{1}{m+1}.$$

In either case,

$$I_m \;=\; \frac{1}{m+1}$$

as asserted.

□

4 Continuous functions, Integrability

The most important bounded functions on a closed interval $[a, b]$ are *continuous functions*. Intuitively, a continuous function is one whose graph one can draw without taking the pen (or whatever else one uses) off the paper. Polynomials and $\sin x$ and $\cos x$ are the functions which immediately spring to mind. There are ways to produce lots of others from simple ones by taking scalar multiples, sums, products, quotients and composites. In practice, even when a function is not continuous, one often wants to approximate it by a continuous one.

It is natural to ask if continuous functions are integrable over $[a, b]$ and amazingly, the answer is *yes*. This will be established in this chapter.

One can also ask about the integrability of bounded functions f on $[a, b]$ which are continuous out side a set S which is *small* in some suitable sense. The simplest case is when S is a finite set of points, and in this case f will be integrable. The right general answer is that the set S of discontinuities of f, i.e, he set of points in $[a, b]$ where f is not continuous must be *negligible* in a suitable sense, i.e, have *measure zero*. This will be discussed at the end of the chapter. If this concept becomes forbidding for some reason, one should try to at least learn that all is well when S is finite.

4.1 Limits and Continuity

We looked at limits of sequences in the Chapter 2. Now we need to go further and consider limits of values of a function.

Let A be a subset of $\mathbb{R}$. Consider any function

$$f : A \to \mathbb{R}$$

At every x in A, $f(x)$ will denote the value of f at x.

Let a be a real number such that f is defined on an open interval around a, except possibly at a itself. Let $L \in \mathbb{R}$. We will say that $f(x)$ **has limit** L **as** x **approaches** a, denoted symbolically by

(4.1.1) $$\lim_{x \to a} f(x) = L,$$

iff the following holds:

For every open interval I centered at L, we can find an open interval I' centered at a such that

(i) $I' - \{a\} \subset A$, and

(ii) $x \in I' - \{a\} \implies f(x) \in I$.

In simple words, we try to approach a from within A, from either direction, and see if $f(x)$ tends to L. In particular, if $\{x_1, x_2, \ldots,\}$ is any sequence in A with limit a, then the

sequence $\{f(x_1), f(x_2), \ldots\}$ converges with limit L. So to prove that some $f(x)$ does not have a limit as x approaches a, it suffices to exhibit a sequence x_n converging to a and show that $f(x_n)$ does not have a limit as n goes to ∞.

Note that to have $\lim_{x\to a} f(x) = L$, it is necessary and sufficient to be able, for every $\varepsilon > 0$, to find some $\delta > 0$ such that

$$|x-a| < \delta \implies |f(x) - L| < \varepsilon.$$

Proposition 4.1.2 *Let $a \in \mathbb{R}$, and let f, g functions defined on an open interval around a, though not necessarily at a itself. Suppose $f(x)$ and $g(x)$ have limits, say L, L' respectively, as x approaches a. Then we have, for all $\alpha, \beta \in \mathbb{R}$,*

$$\lim_{x\to a} (\alpha f + \beta g)(x) \;=\; \alpha L + \beta L',$$

$$\lim_{x\to a} (fg)(x) \;=\; LL',$$

and moreover, if g is non-zero in an open interval around a with $L' \neq 0$,

$$\lim_{x\to a} \frac{f}{g}(x) \;=\; \frac{L}{L'}.$$

The proof is straight-forward and will be left as an exercise. However, the result is very important, and we will use it often.

Example: Let us try to evaluate the following limit:

$$L \;=\; \lim_{x\to 0} \frac{\sqrt{1+x} - \sqrt{1-x}}{x}$$

Since the numerator and the denominator both approach 0 as x goes to 0, we cannot directly apply the Proposition above to express L as the ratio of the limits. How can one resolve this 0/0 problem? The answer is actually simple. One multiplies the numerator and the denominator by $\sqrt{1+x} + \sqrt{1-x}$, and gets

$$L \;=\; \lim_{x\to 0} \frac{(1+x) - (1-x)}{x(\sqrt{1+x} + \sqrt{1-x})} \;=\; \lim_{x\to 0} \frac{2}{\sqrt{1+x} + \sqrt{1-x}}.$$

Evidently, since the numerator is 1 and the denominator has the value 2 at $x = 0$, we see by the Proposition above that L is 1.

The set A where a function f is defined is called the **domain** of f. When a lies in A, f has a value at a and so one can ask for more, and this leads to the notion of **continuity**.

The function f is said to be **continuous at a point** a in A iff we have

$$\lim_{x\to a} f(x) \;=\; f(a)$$

In other words, for f to be continuous at a, we need two things, namely (1) that $f(x)$ has a limit as x approaches a, and (2) that this limit is none other than $f(a)$.

f is said to be a **continuous** function on A iff it is continuous at *every* point a in A.

Simple examples of continuous functions are the *constant function* $f(x) = c$ (for some $c \in \mathbb{R}$), the *linear function* $f(x) = mx + c$, and the *sine* and *cosine* functions.

Another important example is the power function $f(x) = x^n$, for any non-negative integer n. Let us now verify that it is indeed continuous at any point a in $\mathbb{R}$. We may assume that $n > 0$ as $x^0 = 1$ is a constant function, hence continuous. We need to check, for any fixed $n > 0$, that given any $\varepsilon > 0$ one can find a $\delta > 0$ such that

$$-\delta < x - a < \delta \implies -\varepsilon < x^n - a^n < \varepsilon.$$

Now $-\delta < x - a < \delta$ implies that

$$(a - \delta)^n - a^n < x^n - a^n < (a + \delta)^n - a^n.$$

So we have to arrange that

$$(a + \delta)^n - a^n \leq \varepsilon$$

and

$$(a - \delta)^n - a^n \geq -\varepsilon.$$

In other words, we need

$$\delta \leq (a^n + \varepsilon)^{1/n} - a$$

and

$$\delta \leq a - (a^n - \varepsilon)^{1/n}.$$

We can achieve this by taking

$$\delta = \min\{(a^n + \varepsilon)^{1/n} - a, a - (a^n - \varepsilon)^{1/n}\},$$

which is easily seen to be positive. Thus $f(x) = x^n$ is a continuous function for any integer $n \geq 0$. If n is a negative integer, then $f(x) = x^n$ is defined and continuous outside $x = 0$.

Let us next consider the **square root function** $f(x) = \sqrt{x}$, which is continuous at any a in its domain, i.e., for any $a \geq 0$. Let us verify it now. Pick any $\varepsilon > 0$. We have to find a suitable $\delta > 0$ such that

$$(*) \qquad a - \delta < x < a + \delta \implies \sqrt{a} - \varepsilon < \sqrt{x} < \sqrt{a} + \varepsilon.$$

When we write $a - \delta < x$, we have to remember the constraint that $x \geq 0$ for its square-root to make sense. When $a = 0$, we can achieve $(*)$ by taking $\delta = \varepsilon^2$. So assume $a > 0$, and choose δ to be smaller than a. Then

$$a - \delta < x < a + \delta \implies \sqrt{a - \delta} < \sqrt{x} < \sqrt{a + \delta},$$

so we need

$$\sqrt{a+\delta} \;\leq\; \sqrt{a}+\varepsilon$$

and

$$\sqrt{a-\delta} \;\geq\; \sqrt{a}-\varepsilon.$$

This is satisfied by taking

$$\delta \;=\; \min\{(\sqrt{a}+\varepsilon)^2 - a, a - (\sqrt{a}-\varepsilon)^2\}.$$

Done.

More generally, by a similar argument one can prove the following

Proposition 4.1.3 *For any $t \in \mathbb{R}$, the power function $f(x) = x^t$ is continuous in its domain.*

To cook up more continuous functions from the known ones we need another structural result:

Proposition 4.1.4 *Let f, g be continuous functions on a subset A of $\mathbb{R}$. Then any constant multiple cf of f, given by $x \mapsto cf(x)$, the sum function $f+g$, given by $x \to f(x)+g(x)$, and the product function fg, given by $x \mapsto f(x)g(x)$, are all continuous on A. Furthermore, if g is non-zero on A, then the ratio f/g, given by $x \mapsto f(x)/g(x)$, is also continuous.*

In particular, any **polynomial function**

$$f(x) \;=\; a_0 + a_1 x + a_2 x^2 + \ldots + a_n x^n$$

or a function of the form

$$f(x) \;=\; \sqrt{1-x} - \sqrt{x^3 - x + 2}$$

is continuous. So is a **rational function** $f(x)/g(x)$, with f, g polynomials and $g \neq 0$ on A. One knows that $\cos x$ is zero only at the odd integral multiples of $\pi/2$. Hence the **tangent function**, defined by $\tan x = \frac{\sin x}{\cos x}$ is continuous on any subset A of $\mathbb{R}$ avoiding $\{(2k+1)\pi/2 \,|\, k \in \mathbb{Z}\}$.

Suppose we have two functions $g : B \to \mathbb{R}$ and $f : A \to \mathbb{R}$. If A contains the image $g(B)$ of g, then we can define the **composite function**

$$h \;=\; f \circ g \;:\; B \;\to\; \mathbb{R} \tag{4.1.5}$$

by

$$h(x) \;=\; f(g(x)), \quad \forall\, x \in B.$$

Proposition 4.1.6 *If f, g are continuous, then so is their composite $h = f \circ g$, defined as above.*

This allows us to construct a lot of continuous functions, for example, rational functions of trigonometric functions (when defined).

4.2 Some theorems on continuous functions

Intermediate Value Theorem *Let f be a continuous function on a closed interval $[a, b]$. Then for every t between $f(a)$ and $f(b)$ we can find a c in $[a, b]$ such that $f(c) = t$.*

Before proving this result, let us use it to show that there is some real number x such that

$$\sin x \,=\, x - 1.$$

Let us put $f(x) = x - 1 - \sin x$. We have to find x such that $f(x) = 0$. Note that

$$f(0) \,=\, 0 - 1 - 0 \,=\, -1 \,<\, 0,$$

while

$$f(\pi) \,=\, \pi - 1 - \sin\pi \,=\, \pi - 1 \,>\, 0,$$

because $\sin\pi = 0$ and $\pi > 1$. Since $f(x)$ is continuous on $[0, \pi]$, in fact on all of $\mathbb{R}$, we may apply the intermediate value theorem to conclude the existence of some c in $[0, \pi]$ such that $f(c) = 0$. Done.

Proof. There is nothing to prove if t is $f(a)$ or $f(b)$. So we may assume that $f(a) \neq f(b)$. Suppose $f(a) < t < f(b)$. Put $g(x) = f(x) - t$. Then $g(a) = f(a) - t$ is negative and $g(b) = f(b) - t$ is positive. Put

(4.2.1) $$X \,=\, \{u \in [a, b] \mid g(x) < 0 \quad \text{if} \quad a \leq x \leq u\}.$$

Since a lies in X, this set is non-empty, and moreover, b is an upper bound for X as $g(b) > 0$. Put

$$c \,=\, \sup(X),$$

which exists by the properties of $\mathbb{R}$.

We claim that $g(c) = 0$. If $g(c) > 0$, then $g(x)$ will be positive in an open interval centered at c, so that $g(c - \delta) > 0$ for some $\delta > 0$. Since c is the *least* upper bound, any number in (a, c), in particular $c - \frac{\delta}{2}$ will be in X. This gives a contradiction, as $g(c - \delta)$ is positive, violating the definition of X. Hence $g(c)$ is cannot be positive. Suppose $g(c)$ is negative. Then we can find a $\delta > 0$ such that $g(x)$ remains negative in $(c - \delta, c + \delta)$, contradicting the fact that c is an upper bound. So the only possibility is for $g(c)$ to be 0. Then $f(c) = g(c) + t$ will be t, as desired.

We have to still analyze the case when $f(a) > t > f(b)$. Here we apply what we have just proved to $-f$ to conclude what we want.

$\square$

Here is a nice consequence of this result.

Corollary 4.2.2 *Every polynomial $f(x)$ with real coefficients and of* **odd degree** *has a root in $\mathbb{R}$.*

Proof. Let n be the (odd) degree of f. Write

$$f(x) \,=\, a_0 + a_1 x + a_2 x^2 + \ldots a_n x^n,$$

with $a_n \neq 0$. Put

$$g(x) = \frac{1}{a_n} f(x).$$

The coefficients $b_j = \frac{a_j}{a_n}$ are still real and so it suffices to prove that $g(x)$ has a real root.

Intuitively, the argument is clear as the x^n term should dominate all the others when $|x|$ is large, and so $g(x)$ should be positive for large positive x and negative for large negative x. This is what we will now prove rigorously.

Write

$$g(x) = x^n \left(1 + \frac{b_{n-1}}{x} + \ldots + \frac{b_1}{x^{n-1}} + \frac{b_0}{x^n}\right).$$

By applying the triangle inequality repeatedly, we get

$$\left|\frac{b_{n-1}}{x} + \ldots + \frac{b_0}{x^n}\right| \leq \left|\frac{b_{n-1}}{x}\right| + \ldots + \left|\frac{b_0}{x^n}\right|.$$

Now choose x so that

$$|x| \geq \max\{1, 2n|b_{n-j}| \,|\, 1 \leq j \leq n\}. \tag{4.2.3}$$

Then $\frac{|b_{n-j}|}{|x|^j}$ is bounded above by $\frac{1}{2n}$, and

$$\left|\frac{b_{n-1}}{x}\right| + \ldots + \left|\frac{b_0}{x^n}\right| \leq \frac{1}{2}.$$

So we get

$$g(x) = x^n(1 + h(x))$$

with

$$-\frac{1}{2} \leq h(x) \leq \frac{1}{2}$$

if (4.2.3) holds. Now if x is positive (resp.negative) and satisfies (4.2.3), then $g(x)$ will be positive (resp. negative), and we are done by the intermediate value theorem.

□

Theorem 4.2.4 *Let $f : A \to \mathbb{R}$ be a continuous function. Then f takes any closed interval $[a, b]$ contained in A to a compact set in $\mathbb{R}$. In particular, $f([a, b])$ is bounded. Moreover, it attains its* **extremal** *values on $[a, b]$, i.e., there exist c, d in $[a, b]$ such that*

$$f(x) \geq f(c)\,\forall x \in [a, b] \quad \text{and} \quad f(x) \leq f(d)\,\forall x \in [a, b]. \tag{4.2.5}$$

$f(c)$ is called the **minimum** of f on $[a, b]$, and $f(d)$ the **maximum**. The fact that a continuous function attains its extrema on any closed interval is a very important and oft-used fact.

One can generalize further and prove that the image of *any* compact set C under a continuous function f is compact.

When we combine Theorem 4.2.4 with the intermediate value theorem, we get the following useful

Corollary 4.2.6 *Let f be a continuous, monotone increasing function on $[a, b]$. Put*

$$M = \sup\{f(t) \mid t \in [a, b]\} \quad \text{and} \quad m = \inf\{f(t) \mid t \in [a, b]\}. \tag{4.2.7}$$

Then for every t in $[m, M]$ there exists $c \in [a, b]$ such that $f(c) = t$.

Proof of Theorem 4.2.4. Put

$$B = \{x \in [a.b] \mid f([a, x]) \quad \text{bounded}\}.$$

It is bounded, as it is a subset of $[a, b]$, and non-empty because a is in B. So by the properties of $\mathbb{R}$, there is a least upper bound t of B.

Suppose $t < b$. Then the continuity of f at t implies in particular that for every $\varepsilon > 0$, there is some $\delta > 0$ such that

$$x < t + \delta \implies f(x) \in (f(t) - \varepsilon, f(t) + \varepsilon).$$

This implies that f is bounded in $[a, t + \varepsilon/2]$, contradicting the fact that t is the *least* upper bound of B. Hence t cannot be less than b, and we must have $t = b$. So f is bounded on $[a, b]$.

Suppose z is a boundary point of $f([a, b])$. Then we can find a sequence $z_1, z_2, \ldots$ of points lying in $f([a, b])$ converging to z. For each $n \geq 1$ write $z_n = f(y_n)$ with $y_n \in [a, b]$. The sequence $\{y_n\}$ is bounded (as it lies in $[a, b]$) and hence contains a subsequence $\{y_{n_m}\}$which converges, say to $y \in \mathbb{R}$. Such a y must lie in the closure of $[a, b]$, i.e., in $[a, b]$ itself. Then $f(y)$ will, by the continuity of f, be the limit of $\{z_{n_m}\}$, which is a subsequence of $\{z_n\}$. But the sequence $\{z_n\}$ is convergent, and so must have the same limit as the subsequence $\{z_{n_m}\}$. This forces z to be $f(y)$, and thus z lies in $f([a, b])$.

Thus $f([a, b])$ is closed and bounded, i.e., compact.

Now define M, m as in (4.2.7), which makes sense because $f([a, b])$ is bounded. But the supremum M must be a boundary point, for otherwise we can choose a smaller upper bound, contradicting the fact that M is the *least* upper bound. Since $f([a, b])$ is closed, M must belong to $f([a, b])$. Similarly, the infemum m also belongs to $f([a, b])$. Hence there are points c, d in $[a, b]$ such that $m = f(c)$ and $M = f(d)$. Now (4.2.5) is evident.

□

4.3 Integrability of continuous functions

The main result of this section is the following

Theorem 4.3.1 *Every continuous function f on a closed interval $[a, b]$ is integrable.*

The proof rests on a basic property of continuous functions on closed intervals, which we will now expose.

To begin, let us define the **span of f on any closed interval** $[c, d]$ to be

$$\text{span}_f([c, d]) = \max f([c, d]) - \min f([c, d])).$$

The Small Span Theorem *For every $\varepsilon > 0$, there exists a partition $P: a = t_0 < t_1 < \ldots < t_n = b$ of $[a, b]$ such that* $\text{span}_f([t_{j-1}, t_j]) < \varepsilon$, for each $j = 1, 2, \ldots, n$.

Proof. We will prove this by contradiction. Suppose the theorem is false. Then there exists some $\varepsilon > 0$, call it ε_0 to indicate that it is fixed, such that for **every** partition $P: a = t_0 < t_1 < \ldots < t_n = b$, $\text{span}_f([t_{j-1}, t_j]) \geq \varepsilon_0$ for some j. Subdivide $[a, b]$ into two closed intervals $[a, c], [c, b]$ with c being the midpoint $(a + b)/2$. Then the theorem must be false for one of these subintervals, call it J_1, for the **same** ε_0. Do this again and again, and we finally end up with an infinite sequence of **nested** closed intervals

$$[a, b] = R_0 \supset R_1 \supset R_2 \supset \ldots,$$

such that, for every $m \geq 0$, the span of f is at least ε_0 for **any** partition $P_m = \{J_{j,m}\}$ of R_m on some $J_{j,m}$. Let x_m denote the left endpoint of R_m, for each $m \geq 0$. Then the sequence $\{x_m\}_{m \geq 0}$ is bounded, and so we may find the least upper bound (**sup**) γ, say, of x_m. Then γ will be a boundary point of $[a, b]$, and hence must lie in it, as $[a, b]$ is closed. Since f is continuous at γ, we can find a closed subinterval I of $[a, b]$ containing γ such that $\text{span}_f(S) < \varepsilon_0$. But by construction R_m will have to lie inside I if m is large enough, say for $m \geq M$. This gives a contradiction to the span of f being $\geq \varepsilon_0$ on some open set of every partition of R_M. Thus the small span theorem holds for $(f, [a, b])$.

$\square$

Proof of Theorem 4.3.2. Recall that it suffices to show that, given any $\varepsilon > 0$, there is a partition P of $[a, b]$ such that $U(f, P) - L(f, P) < \varepsilon$. Put $\varepsilon_1 = \varepsilon/(b - a)$. Applying the small span theorem, we can find a partition $P = \{J_j\}$ such that $\text{span}_f(J_j) < \varepsilon_1$, for all j. Then clearly,

$$U(f, P) - L(f, P) < \varepsilon_1 (b - a) = \varepsilon.$$

Done.

$\square$

4.4 Trigonometric functions

We will assume, as we have all along, that the students taking this class are familiar with basic trigonometric functions, like the $\sin x$, $\cos x$, then $\tan x$, which is defined to be $\frac{\sin x}{\cos x}$ whenever $\cos x$ is non-zero, as well as the functions

$$\csc x = \frac{1}{\sin x}, \quad \sec x = \frac{1}{\cos x}, \quad \text{and} \quad \cot x = \frac{1}{\tan x} = \frac{\cos x}{\sin x}.$$

We will accept the following from **Basic Trigonometry** without proof:

Theorem 4.4.1

(a) $\sin x$ *and* $\cos x$ *are periodic functions with* **period** 2π*, i.e.,*

$$\sin(x \pm 2\pi) = \sin x \quad \text{and} \quad \cos(x \pm 2\pi) = \cos x,$$

with

$$\sin^2 x + \cos^2 x = 1.$$

(b) $\sin x$ *is an* **odd** *function, while* $\cos x$ *is an* **even** *function, i.e.,*

$$\sin(-x) = -\sin x \quad \text{and} \quad \cos(-x) = \cos x.$$

(c) $\sin x$ *and* $\cos x$ *are both non-negative, monotone functions in* $[0, \pi/2]$*, with* $\sin x$ *increasing (from* 0 *to* 1*) and* $\cos x$ *decreasing (from* 1 *to* 0*) as* x *moves from* 0 *to* $\pi/2$*.*

(d) *(***Addition theorems)**

$$\sin(x \pm y) = \sin x \cos y \pm \cos x \sin y$$

and

$$\cos(x \pm y) = \cos x \cos y - \pm \sin x \sin y.$$

(e) *For every* x *in the open interval* $(0, \pi/4)$*,*

$$\cos x < \frac{\sin x}{x} < 1.$$

As a consequence of (e), we can compute some important limits. For example, we obtain

Proposition 4.4.2

(i)
$$\lim_{x \to 0} \frac{\sin x}{x} = 1,$$

(ii)
$$\lim_{x \to 0} \frac{1 - \cos x}{x} = 0.$$

Proof. For all small x, $\frac{\sin x}{x}$ is caught, by part (e) of Theorem 4.4.1, between the constant value 1 and $\cos x$. But $\cos x$ becomes 1 as x approaches 0. This gives (i).

To get (ii), multiply the numerator $(1 - \cos x)$ and the denominator x by $1 + \cos x$ and note that

$$(1 - \cos x)(1 + \cos x) = 1 - \cos^2 x = \sin^2 x.$$

So we get, by applying Proposition 4.1.2,

$$\lim_{x \to 0} \frac{1 - \cos x}{x} = \left(\lim_{x \to 0} \frac{\sin x}{x} \right) \left(\lim_{x \to 0} \frac{\sin x}{1 + \cos x} \right) = 1 \cdot 0 = 0.$$

Here we have used (i) as well as the facts that $\sin 0 = 0$ and $1 + \cos 0 = 2$.

$\square$

The following result is basic.

Theorem 4.4.3 *Let* $[a, b]$ *be any closed interval. Then the functions* $\sin x$ *and* $\cos x$ *are integrable on* $[a, b]$. *Explicitly,*

$$\int_a^b \sin x dx = \cos a - \cos b$$

and

$$\int_a^b \cos x dx = \sin b - \sin a.$$

Proof. By periodicity (part (a) of Theorem 4.4.1) and additivity of the integral (Proposition 3.x), we may assume that $0 \le a < b \le 2\pi$. Moreover, for $0 \le x \le \pi$, the *oddness* of $\sin x$ implies

$$\sin(\pi + x) = \sin(-\pi + x) = -\sin(\pi - x), \tag{4.4.4}$$

while the *evenness* of $\cos x$ implies

$$\cos(\pi + x) = \cos(-\pi + x) = \cos(\pi - x).$$

Moreover, the trigonometric definition of $\sin x$ and $\cos x$ gives immediately the following identities:

$$\sin(\pi - x) = \sin x \tag{4.4.5}$$

and

$$\cos(\pi - x) = -\cos x.$$

Thanks to (4.4.4) and (4.4.5) it suffices to prove the assertion of the Theorem when $0 \le a < b \le \pi/2$. Furthermore, since

$$\int_a^b f(x)dx = \int_0^b f(x)dx - \int_0^a f(x)dx$$

for any function $f(x)$, it suffices to prove that for any a in $(0, \pi/2]$, the functions $\sin x$ and $\cos x$ are integrable on $[0, a]$ and that

$$\int_0^a \sin x\, dx = 1 - \cos a \quad \text{and} \quad \int_0^a \cos x\, dx = \sin a. \tag{4.4.6}$$

We will prove the formula for the integral of $\sin x$ over $[0, a]$ and leave the proof of the corresponding one for $\cos x$ to the reader.

For every $n \geq 1$, define a partition P_n of $[0, a]$ to be given by

$$0 = t_0 \,<\, t_1 = \frac{a}{n} \,<\, t_2 = \frac{2a}{n} \,<\, \ldots < t_n = a, \tag{4.4.7}$$

so that $t_{j+1} - t_j = \frac{a}{n}$ for all $j \leq n-1$.

Since $\sin x$ is a monotone increasing function in $[0, \pi/2]$, the upper and lower sums are given by

$$U(\sin x, P_n) \;=\; \frac{a}{n} \sum_{j=0}^{n-1} \sin(\frac{(j+1)a}{n}) \tag{4.4.8}$$

and

$$L(\sin x, P_n) \;=\; \frac{a}{n} \sum_{j=0}^{n-1} \sin(\frac{ja}{n}).$$

By the addition theorem for $\cos x$ (part (d) of Theorem 4.4.1), we have

$$2 \sin x \sin y \;=\; \cos(x-y) - \cos(x+y). \tag{4.4.9}$$

In particular, when $y = \frac{a}{2n}$ and $x = \frac{(j+1)a}{n}$, we get

$$2 \sin(\frac{a}{n}) \sin(\frac{(j+1)a}{n}) \;=\; \cos(\frac{(2j+1)a}{2n}) - \cos(\frac{(2j+3)a}{2n}). \tag{4.4.10}$$

Applying this in conjunction with (4.4.8), we get

$$U(\sin x, P_n) \;=\; \frac{a/2n}{\sin(\frac{a}{2n})} \left((\cos(\frac{a}{2n}) - \cos(\frac{3a}{2n})) + \ldots + (\cos(\frac{(2n-1)a}{2n}) - \cos(\frac{(2n+1)a}{2n})) \right),$$

which simplifies to

$$U(\sin x, P_n) \;=\; \frac{a}{2n \sin(\frac{a}{2n})} \left(\cos(\frac{a}{2n}) - \cos(\frac{(2n+1)a}{2n}) \right). \tag{4.4.11}$$

By Proposition 4.4.2,

$$\lim_{n \to \infty} \frac{\sin(\frac{a}{2n})}{\frac{a}{2n}} \;=\; 1. \tag{4.4.12}$$

Also, since $\cos 0 = 1$,

$$\lim_{n \to \infty} \cos(\frac{a}{2n}) - \cos(\frac{(2n+1)a}{2n}) \;=\; 1 - \cos a. \tag{4.4.13}$$

Combining (4.4.11) through (4.4.13), we get

(4.4.14)
$$\lim_{n\to\infty} U(\sin x, P_n) \;=\; 1 - \cos a.$$

The analog of (4.4.11) for the n-th lower sum is

(4.4.15)
$$L(\sin x, P_n) \;=\; \frac{a}{2n\sin(\frac{a}{2n})}\left(1) - \cos(\frac{(2n-1)a}{2n})\right).$$

We get the same limit, namely

(4.4.16)
$$\lim_{n\to\infty} L(\sin x, P_n) \;=\; 1 - \cos a.$$

In view of (4.4.14), (4.4.15) and Lemma 3.2.8, the desired identity (4.4.6) follows for the sine integral. The argument is entirely analogous for the cosine integral.

□

4.5 Functions with discontinuities

One is very often interested in being able to integrate bounded functions over $[a, b]$ which are continuous except on a subset which is very small, for example outside a **finite set**. To be precise, we say that a subset Y of $\mathbb{R}$ is **negligible**, or that it has **measure zero**, iff for every $\varepsilon > 0$, we can find a countable number of closed intervals $J_1, J_2, \ldots$ such that
(4.5.1)

(i) $Y \subset \cup_{i=1}^{\infty} J_i$, and

(ii) $\sum_{i=1}^{\infty} \ell(J_i) < \varepsilon$.

If we can do this with a *finite* number of closed intervals $\{J_i\}$ (for each ε), then we will say that Y has **content zero**.

Examples:
(4.5.2)

(1) Any finite set of points in $\mathbb{R}$ has content zero. (Proof is obvious!)

(2) Any subset Y of $\mathbb{R}$ which contains a non-empty open interval (a, b) is **not** negligible.
Proof. It suffices to prove that (a, b) has non-zero measure for $a < b$ in $\mathbb{R}$. Suppose (a, b) is covered by a finite union of a countable collection of closed intervals $J_1, J_2, \ldots$ in $\mathbb{R}$. Then clearly,

$$S := \sum_{i=1}^{m} \text{length}(J_i) \;\geq\; \ell(a, b) \;=\; b - a.$$

So we can never make S less than $b - a$.

Theorem 4.5.3 *Let f be a bounded function on $[a, b]$ which is continuous except on a subset Y of measure zero. Then f is integrable on $[a, b]$.*

Proof. Let $M > 0$ be such that $|f(x)| \leq M$, for all $x \in [a, b]$. Since Y has content zero, we can find closed subintervals $J_1, J_2, \ldots$ of $[a, b]$ such that

(i) $Y \subseteq \cup_{i=1}^{m} J_i$, and

(ii) $\sum_{i=1}^{\infty} \ell(J_i) < \frac{\varepsilon}{4M}$.

Extend $\{J_1, \ldots, J_m\}$ to a partition $P = \{J_1, \ldots, J_r\}$, $m < r$, of $[a, b]$. Applying the small span theorem, we may suppose that $J_{m+1} \ldots, J_r$ are so chosen that (for each $i \geq m + 1$) $\text{span}_f(J_i) < \frac{\varepsilon}{2\ell([a,b])}$. (We can apply this theorem because f is continuous outside the union of $J_1, \ldots, J_m$.) So we have

$$U(f, P) - L(f, P) \;\leq\; 2M \sum_{i=1}^{m} \ell(J_i) + \sum_{i=m+1}^{r} \text{span}_f(J_i)\ell(J_i),$$

which is

$$<\; (2M)\left(\frac{\varepsilon}{2M}\right) + \frac{\varepsilon}{2\ell([a,b])} \sum_{i=m+1}^{r} \ell(J_i) \;\leq\; \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,$$

because $\sum_{i=m+1}^{r} \ell(J_i) \leq b - a$.

□

Remark 4.5.4 We can use this theorem to define the integral of a **continuous function** f **on any compact set** B in $\mathbb{R}$ **if the boundary** of B is **negligible**. Indeed, in such a case, we may enclose B in a closed interval $[a, b]$ and define a function $\tilde{f}$ on $[a, b]$ by making it equal f on B and 0 on $[a, b] - B$. Then $\tilde{f}$ will be continuous on all of $[a, b]$ except on the boundary of B, which has content zero. So $\tilde{f}$ is integrable on $[a, b]$. Since $\tilde{f}$ is 0 outside B, it is reasonable to set

$$\int_B f = \int_a^b \tilde{f}. \tag{4.5.5}$$

5 Improper Integrals, Areas, Polar Coordinates, Volumes

We have now attained a good level of understanding of integration of nice functions f over closed intervals $[a, b]$. In practice one often wants to extend the domain of integration and consider unbounded intervals such as $[a, \infty)$ and $(-\infty, b]$. The simplest non-trivial examples are the **infinite trumpets** defined by the areas under the graphs of x^t for $t > 0$, i.e., the *improper integrals*

$$A_t = \int_1^\infty \frac{1}{x^t} dx.$$

We will see below that A_t has a well defined meaning if $t > 1$, but becomes unbounded for $t \leq 1$.

One is also interested in integrals of functions f over finite intervals with f being unbounded. The natural examples are given (for $t > 0$) by

$$B_t = \int_0^1 \frac{1}{x^t} dx.$$

Here it turns out that B_t is well defined, i.e., has a finite value, if and only if $t < 1$. In particular, neither A_1 nor B_1 makes sense. Moreover,

$$A_t + B_t = \int_0^\infty \frac{1}{x^t} dx$$

is unbounded for every $t > 0$.

After we learn the fundamental theorems of Calculus, we will visit the land of improper integrals again with new tools.

5.1 Improper Integrals

Let f be a function defined on the interior of a possibly infinite interval J such that either its upper endpoint – call it b, is ∞ or f becomes unbounded as one approaches b. But suppose that the lower endpoint – call it a, is finite and that $f(a)$ is defined. In the former case the interval is unbounded, while in the latter case the interval is bounded, but the function is unbounded. We will say that the **integral of f over J** exists iff the following two conditions hold:
(5.1.1)

(i) For every $u \in (a, b)$, f is integrable on $[a, u]$; and

(ii) the limit

$$\lim_{u\to b,\, u<b} \int_a^u f(x)dx$$

exists.

When this limit exists, we will call it the **integral of f over J** and write it symbolically as

$$\int_a^b f(x)dx.$$

Similarly, if a is either $-\infty$ or is a finite point where f becomes unbounded, but with b a finite point where f is defined, one sets

$$\int_a^b f(x)dx = \lim_{u\to a,\, u>a} \int_u^b f(x)dx \tag{5.1.2}$$

when the limit on the right makes sense.

Lemma 5.1.3 *We have*

$$\int_a^\infty f(x)dx = \int_a^c f(x)dx + \int_c^\infty f(x)dx \quad \forall\, c \in (a,\infty)$$

and

$$\int_{-\infty}^b f(x)dx = \int_{-\infty}^c f(x)dx + \int_c^b f(x)dx \quad \forall\, c \in (-\infty, b),$$

when the integrals make sense.

Proof. We will prove the first **additivity formula** and leave the other as an exercise for the reader. Pick any $c \in (a,\infty)$. or all real numbers $u > c$, we have by the usual additivity formula,

$$\int_a^u f(x)dx = \int_a^c f(x)dx + \int_c^u f(x)dx.$$

Thus

$$\int_a^\infty f(x)dx = \lim_{u\to\infty} \left(\int_a^c f(x)dx + \int_c^u f(x)dx \right),$$

which equals

$$\int_a^c f(x)dx + \lim_{u\to\infty} \int_c^u f(x)dx = \int_a^c f(x)dx + \int_c^\infty f(x)dx.$$

$\square$

If J is an interval with both of its endpoints being problematic, we will choose a point c in (a, b) and put

$$\int_a^b f(x)dx = \int_a^c f(x)dx + \int_c^b f(x)dx \tag{5.1.4}$$

if both the improper integrals on the right make sense. One can check using Lemma 5.1.3 above that this definition is independent of the choice of c.

When an improper integral does not make sense, we will call it **divergent**. Otherwise it is **convergent**.

Proposition 5.1.4 *Let t be a positive real number. Then for $t > 1$,*

$$\int_1^\infty \frac{1}{x^t}dx = \frac{1}{t-1}, \tag{A}$$

with the improper integral on the left being divergent for $t \leq 1$.

On the other hand, if $t \in (0, 1)$,

$$\int_0^1 \frac{1}{x^t}dx = \frac{1}{1-t}, \tag{B}$$

with the improper integral on the left being divergent for $t \geq 1$.

Proof. We learnt in chapter 3 that for all $a, b \in \mathbb{R}$ with $a < b$, and for all $t \neq 1$:

$$\int_a^b \frac{1}{x^t}dx = \frac{b^{1-t} - a^{1-t}}{1-t}.$$

Hence for $t > 1$,

$$\lim_{u\to\infty} \int_1^u \frac{1}{x^t}dx = \lim_{u\to\infty} \frac{1}{(1-t)u^{t-1}} + \frac{1}{t-1} = \frac{1}{t-1},$$

because the term $\frac{1}{u^{t-1}}$ goes to zero as u goes to ∞. If $t < 1$, this term goes to ∞ as u goes to ∞, and so the integral is divergent. Finally let $t = 1$. Then we cannot use the above formula. But for each $N \geq 1$, we have the inequality

$$\sum_{n=1}^N \frac{1}{n} \leq \int_1^N \frac{1}{x}dx.$$

The reason is that the sum on the left is a lower Riemann sum for the function $f(x) = \frac{1}{x}$ over the interval $[1, N]$ relative to the partition $P : 1 < 2 < \ldots < N$. So, if the improper integral of this function exists over $[1, \infty)$, the infinite series

$$\sum_{n=1}^{\infty} \frac{1}{n} = \lim_{N \to \infty} \sum_{n=1}^{N} \frac{1}{n}$$

must converge. But we have seen in chapter 2 that this series diverges. So the integral is divergent for $t = 1$, and (A) is proved in all cases.

The proof of (B) is very similar and will be left as an exercise.

□

5.2 Areas

Given any region R in the plane, one is often interested in assigning, if possible, an **area**. Whenever the area makes sense, R will be called a **measurable region**. If the region is bounded, one can proceed in the following intuitive way. For each $n > 1$, enclose R in a rectangular grid made up of squares of side $1/n$. Add up the areas of all the squares in the grid which meet R, and denote this (upper) sum by $U_n(R)$. Then add up the areas of all the squares which are contained in R, and call this (lower) sum $L_n(R)$. Then let n become very large and see if the two sums converge to the same limit. In the event they do, we will say that R is **measurable** and call the common limit the **area of** R, denoted by $A(R)$. Of course this procedure should be reminiscent of the way we evaluated the integral of a nice function f on a closed interval $[a, b]$. Indeed, the definite integral was defined in such a way that when it exists and when f is non-negative on $[a, b]$, its value will equal the area under the graph of f over $[a, b]$. To be precise we should say that the value is the area of the region caught between the graph of f over $[a, b]$, the x-axis and the two vertical lines $x = a$ and $x = b$.

Now that we have defined improper integrals, we can also consider **areas of infinite plane regions** and define them by a **limiting process**. But we should note that to get a finite answer, the region will need to get narrower and narrower as the x-coordinate gets near ∞ or $-\infty$. Similarly, we can consider regions of bounded width, but unbounded vertically; for the area to exist here, the region will need to get very narrow as the y-coordinate gets near ∞ or $-\infty$.

Here are some basic properties of the area function of measurable sets in the plane:
(5.2.1)

(i) $A(R) \geq 0$.

(ii) $A(R \cup R') = A(R) + A(R') - A(R \cap R')$.

(ii) $A(R)$ does not change if R gets translated or rotated in the plane.

Proposition 5.2.2 *Let J be an interval, possibly of infinite length. Let f, g be functions which are defined on the interior of J and are integrable on J. Denote by R the region between the graph of f, graph of g, and the vertical lines at the end points of J unless J is of infinite length, in which case one takes the vertical lines only at the finite endpoint if any. Then R is measurable and*

$$A(R) = \int_J |f-g|(x)dx,$$

where $|f-g|$ denote the function $x \mapsto |f(x)-g(x)|$.

For example let us calculate $A(R)$ when $J = [0, \pi/2]$, $f(x) = \cos x$ and $g = \sin x$. Then $f - g$ is non-negative in $[0, \pi/4]$ and non-positive in $[\pi/4, \pi/2]$. So we get, by the additivity of the integral,

$$A(R) = \int_0^{\pi/4} (\cos x - \sin x)dx + \int_{\pi/4}^{\pi/2} (\sin x - \cos x)dx.$$

Recall that for all a, b with $a < b$,

$$\int_a^b \sin x dx = \cos a - \cos b$$

and

$$\int_a^b \cos x dx = \sin b - \sin a.$$

So by the linearity of the integral,

$$A(R) = \{(\sin \pi/4 - \sin 0) - (\cos 0 - \cos \pi/4)\} + \{(\cos \pi/4 - \cos \pi/2) - (\sin \pi/2 - \sin \pi/4)\}.$$

One knows that $\sin 0 = \cos \pi/2 = 0$, $\cos 0 = \sin \pi/2 = 1$, and $\sin \pi/4 = \cos \pi/4 = 1/\sqrt{2} = \sqrt{2}/2$. Hence

$$A(R) = \frac{\sqrt{2}}{2} - 0 - 1 + \frac{\sqrt{2}}{2} + \frac{\sqrt{2}}{2} - 0 - 1 + \frac{\sqrt{2}}{2} = 2(\sqrt{2} - 1).$$

Sometimes one speaks of the area of the region R between the graphs of two functions without specifying the base interval. In that case R is taken to be the bounded region, if any, caught between the graphs. If the graphs do not meet or if they meet just once, then there is no bounded region.

For *example*, consider the functions

$$f(x) = x^2 \quad \text{and} \quad g(x) = x^3 - x.$$

To find the bounded region R caught between the graphs, one has to first solve for the points (if any) where the graphs meet. This is the same as finding the points x where $x^2 = x^3 - x$,

i.e., $x = 0$ or $x^2 - x - 1 = 0$. There are three real solutions, namely $x = 0, \frac{1+\sqrt{5}}{2}$ and $\frac{1-\sqrt{5}}{2}$. Note that $x^3 - x \geq x^2$ on $[\frac{1-\sqrt{5}}{2}, 0]$, and $x^3 - x \leq x^2$ on $[0, \frac{1+\sqrt{5}}{2}]$. (Sketch the graphs and see!) So we get

$$A(R) = \int\limits_{(1-\sqrt{5})/2}^{0} (x^3 - x - x^2)dx - \int_0^{(1+\sqrt{5})/2} (x^3 - x - x^2)dx.$$

Since $\int\limits_a^b x^n dx$ is, for $n \neq -1$, $(b^{n+1} - a^{n+1})/(n+1)$,

$$A(R) = -\frac{(1-\sqrt{5})^4}{64} + \frac{(1-\sqrt{5})^2}{8} + \frac{(1-\sqrt{5})^3}{24} + \frac{(1+\sqrt{5})^4}{64} - \frac{(1+\sqrt{5})^2}{8} - \frac{(1+\sqrt{5})^3}{24}.$$

This simplifies to give

$$A(R) = \frac{5\sqrt{5}}{12}.$$

5.3 Polar coordinates

So far we have confined ourselves to **rectilinear coordinates** on the plane, which are often called **Cartesian coordinates** to honor René Descartes who introduced them. Simply put, we identify each point P on the plane by the pair (x, y), where x (resp. y) is the distance between the origin O and the point where the x-axis (resp. y-axis) meets the perpendicular to it from P. Instead one can look at the pair (r, θ), where r is the distance from P to the origin, measured on the line L connecting O to P, and the angle between the x-axis and L, measured in the counterclockwise direction. By definition $r \geq 0$, and we will take θ to lie in $[0, 2\pi)$. In particular, the angle is taken to be zero, and not 2π or 4π or -2π, for any point lying on the x-axis. It should also be noted that as defined, the angle does not make much sense for the origin; we take (r, θ) to be $(0, 0)$ for it.

The quantities r, θ are called the **polar coordinates** of P. It is easy to see that

$$x = r\cos\theta \quad \text{and} \quad y = r\sin\theta. \tag{5.3.1}$$

In the reverse direction, one can almost recover (r, θ) from (x, y) by the easily verified formulae

$$r = \sqrt{x^2 + y^2} \quad \text{and} \quad \tan\theta = \frac{y}{x}. \tag{5.3.2}$$

We said *almost*, because we do not know at the moment to what extent $\tan\theta$ determines θ. We will come back to this question after studying inverse functions.

Sometimes the equation defining a curve, or a region, in the plane becomes simpler when we use polar coordinates, and this is the reason for studying them. For example, the circle

centered at O defined by the equation $x^2 + y^2 = a^2$ can be easily described as the graph of $r = a$. The region inside the circle is simply $r \leq a$. The rule of thumb is that whenever there is **circular symmetry** in a given situation, it is better to use polar coordinates.

Suppose S is an **angular sector**, i.e., the region bounded by $\theta = a, \theta = b$ and $r = \rho$, with $b - a \in [0, 2\pi]$. ρ is called the **radius**, and $b - a$ the **angle**, of s. Using the definition of π as the area of the region inside the unit circle, one can show

$$A(S) = \frac{1}{2}(b-a)\rho^2. \tag{5.3.3}$$

We will accept this as a basic fact.

The main problem here will be to understand the **radial sets**, and to know when they are measurable. Such a set is given as the region bounded by $\theta = a, \theta = b$ and $r = f(\theta)$, where f is a function of $[a, b]$ and $b - a \in [0, 2\pi]$. One can use Calculus to find the area of R under a hypothesis on f.

Let us call a function f on $[a, b]$ **square-integrable** iff f^2 is integrable on $[a, b]$. Note that every continuous function on $[a, b]$ is square-integrable.

Proposition 5.3.4 *Let R be a radial set bounded by $\theta = a, \theta = b$ and $r = f(\theta)$, where f is a square-integrable function of $[a, b]$ and $b - a \in [0, 2\pi]$. Then*

$$A(R) = \frac{1}{2}\int_a^b f(\theta)d\theta.$$

Proof. Pick any partition

$$P: \, a = t_0 < t_1 < \ldots < t_n = b.$$

For every integer j with $1 \leq j \leq n$, write

$$m_j = \inf f([t_{j-1}, t_j]) \quad \text{and} \quad M_j = \sup f([t_{j-1}, t_j]), \tag{5.3.5}$$

and denote by S_j, resp. S'_j, the angular sector of radius m_j, resp. M_j, between $\theta = t_{j-1}$ and $\theta = t_j$. Then

$$\frac{1}{2}L(f^2, P) = \frac{1}{2}\sum_{j=1}^{n}(T_j - t_{j-1})m_j^2 = \sum_{j=1}^{n} A(S_j) \tag{5.3.6}$$

and

$$\frac{1}{2}U(f^2, P) = \frac{1}{2}\sum_{j=1}^{n}(T_j - t_{j-1})M_j^2 = \sum_{j=1}^{n} A(S'_j).$$

It follows that

$$\frac{1}{2}L(f^2, P) \leq A(R) \leq \frac{1}{2}U(f^2, P). \tag{5.3.7}$$

Since f^2 is integrable on $[a, b]$ by hypothesis, the upper sums and lower sums (of f^2) converge to a common limit, which is the integral of f^2 over $[a, b]$. Now the assertion of the Proposition follows by virtue of (5.3.7).

$\square$

As an **example**, let us look at the region R bounded by the **spiral of Archimedes**:

$$0 \leq \theta \leq 2\pi, \ r = \theta. \tag{5.3.8}$$

Since $f(\theta) = \theta$ is square-integrable we may apply Proposition 5.3.4 and deduce that

$$A(R) = \frac{1}{2}\int_0^{2\pi} \theta^2 d\theta = \frac{1}{2}\frac{(2\pi)^3}{3} = \frac{4\pi^3}{3}.$$

5.4 Volumes

Given a three-dimensional shape V, a basic question which arises is whether we could assign a **volume** to it. When V is bounded one could give an intuitive definition, much like we did for the area of a bounded surface. For every $n > 1$, enclose V in a *three-dimensional grid* consisting of cubes whose sides are of length $1/n$. Define the upper sum U_n, resp. the lower sum L_n, to be the sum of the volumes of all the cubes which meet V, resp. all the cubes which are enclosed in V. As n increases, one gets a further refinement of the grid. We will say that V has a volume, or that V is a **measurable subset of three-space** iff the sequences $\{U_n\}$ and $\{L_n\}$ both converge and have a common limit, which will be denoted as $\mathrm{vol}(V)$.

It is not hard to see that V is itself a **cube**, then its volume is a^3 if a is the length of any **edge** of V. It can also be shown that if V has **constant thickness**, then its volume is the area of the cross-section times the thickness. In particular, the volume of any **cylinder** V of thickness h and **circular cross-section** of radius ρ, $\mathrm{vol}(V)$ is $\pi\rho^2 h$. Going one step further, if V is a **cylindrical shell** of thickness h and **annular cross-section** of **outer radius** ρ_1 and **inner radius** ρ_2, then $\mathrm{vol}(V)$ is $\pi(\rho_1^2 - \rho_2^2)h$. We will need to use these facts below.

The three-dimensional shapes which we can try to understand right now using just one-variable Calculus are the ones obtained by revolving about either axis a plane region under the graph of a (one-variable) function.

Proposition 5.4.1 *Let f be a non-negative, square-integrable function on a closed interval $[a, b]$, and let R be the region under the graph of f between $x = a$ and $x = b$. Denote by V the three-dimensional space obtained by revolving R about the x-axis. Then V is a measurable subset of three-space and moreover,*

$$\mathrm{vol}(V) = \pi \int_a^b f(x)^2 dx.$$

Proof. Choose any partition

$$P: \, a = t_0 < t_1 < \ldots < t_n = b,$$

and for every integer j between 1 and n, let m_j, M_j be as in (5.3.5). We get two cylinders of thickness $t_j - t_{j-1}$ and circular cross-sections of radii m_j, M_j. Clearly we have

$$\pi \sum_{j=1}^{n} (t_j - t_{j-1}) m_j^2 \, \leq \, \mathrm{vol}(V) \, \leq \, \pi \sum_{j=1}^{n} (t_j - t_{j-1}) M_j^2.$$

The expression on the left (resp. right) is just the lower (resp. upper) sum $L(\pi f^2, P)$ (resp. $U(\pi f^2, P)$).

Since f^2 is by assumption integrable on $[a, b]$, the least upper bound $\underline{I}(\pi f^2)$ of $\{L(\pi f^2, P)\}$ equals the greatest lower bound $\overline{I}(\pi f^2)$ of $\{U(\pi f^2, P)\}$. On the other hand, the bounds above imply that

$$\underline{I}(\pi f^2) \, \leq \, \mathrm{vol}(V) \, \leq \overline{I}(\pi f^2).$$

The assertion now follows.

□

When we revolve such an R about the y-axis, it becomes even more interesting.

Proposition 5.4.2 *Let f be a function on a closed interval $[a, b]$ with $a \geq 0$, and let R be the region caught between the graph of f, the x-axis and the vertical lines $x = a$ and $x = b$. Denote by W the three-dimensional space obtained by revolving R about the y-axis. If the function $x \mapsto xf(x)$ is integrable on $[a, b]$, then W is a measurable subset of three-space and moreover,*

$$\mathrm{vol}(W) \, = \, 2\pi \int_a^b xf(x)dx.$$

Proof. For any partition

$$P: \, a = t_0 < t_1 < \ldots < t_n = b,$$

and for every integer j between 1 and n, we get two cylindrical shells of thickness m_j, M_j and annular cross-section of inner radius t_{j-1} and outer radius t_j. Then we get

$$\pi \sum_{j=1}^{n} (t_j^2 - t_{j-1}^2) m_j \, \leq \, \mathrm{vol}(W) \, \leq \, \pi \sum_{j=1}^{n} (t_j^2 - t_{j-1}^2) M_j.$$

Since $t_j^2 - t_{j-1}^2 = (t_j - t_{j-1})(t_j + t_{j-1})$, we get

$$\pi(A_1(P) + A_2(P)) \, \leq \, \mathrm{vol}(W) \, \leq \, \pi(A_3(P) + A_4(P)), \tag{5.4.3}$$

where

$$A_1(P) \, = \, \sum_{j=1}^{n} (t_j - t_{j-1}) t_j m_j,$$

$$A_2(P) \;=\; \sum_{j=1}^{n} (t_j - t_{j-1}) t_{j-1} m_j,$$

$$A_3(P) \;=\; \sum_{j=1}^{n} (t_j - t_{j-1}) t_j M_j,$$

and

$$A_4(P) \;=\; \sum_{j=1}^{n} (t_j - t_{j-1}) t_{j-1} M_j.$$

We now **claim** that for every $\varepsilon > 0$, we can find a partition P such that for each $i \leq 4$,

$$|\int_a^b xf(x)dx \;-\; A_i| \;<\; \varepsilon.$$

This is clear for $i = 2, 4$, because for any P, A_2 (resp. A_4) is $L(g, P)$ (resp. upper sum $U(g, P)$), where g is the function $x \mapsto xf(x)$. But then if we make the partition very fine, making each $t_j - t_{j-1}$ very small, we can make $A_1(P) - A_2(P)$ and $A_3(P) - A_4(P)$ as small as we want. The claim follows.

Consequently, for every $\varepsilon > 0$, there exists P such that

$$|2\int_a^b xf(x)dx \;-\; A_1(P) + A_2(P)| \;<\; 2\varepsilon \tag{5.4.4}$$

and

$$|2\int_a^b xf(x)dx \;-\; A_1(P) + A_2(P)| \;<\; 2\varepsilon.$$

The assertion of the Proposition now follows by combining (5.4.3) and (5.4.4).

□

For a worked example, see problem (E) Assignment 5, for which a solution set has been posted on the web.

5.5 The integral test for infinite series

When we discussed the question of convergence of infinite series in chapter 2, we gave various tests one could use for this purpose, at least for series with non-negative coefficients. Here is another test, which can at times be helpful.

Proposition 5.5.1 *Consider an infinite series*

$$S \;=\; \sum_{n=1}^{\infty} a_n,$$

whose coefficients satisfy

$$a_n \;=\; f(n),$$

for some non-negative, monotone decreasing function f on the infinite interval $[1, \infty)$. Then S converges iff the improper integral

$$I \;=\; \int_1^\infty f(x)dx$$

converges.

Proof. For any integer $N > 1$, consider the partition

$$P_N : \, 1 < 2 < \ldots < N$$

of the closed interval $[0, N]$. Then, since f is monotone decreasing, the upper and lower sums are given by

$$U(f, P_N) \;=\; a_1 + a_2 + \ldots + a_{N-1}$$

and

$$L(f, P_N) \;=\; a_2 + \ldots + a_{N-1} + a_N.$$

Suppose f is integrable over $[1, \infty)$. Then it is integrable over $[1, N]$ and

$$a_1 + a_2 + \ldots + a_{N-1} \;\leq\; \int_1^N f(x)dx \;\leq\; a_2 + \ldots + a_{N-1} + a_N.$$

As N goes to infinity, this gives

$$S \;\leq\; \int_1^\infty f(x)dx,$$

which implies that S is convergent.

To prove the converse we need to be a bit more wily. Suppose S converges. Note that f is integrable over $[1, \infty)$ iff the series

$$T \;=\; \sum_{n=1}^\infty b_n$$

converges, where

$$b_n \;=\; \int_n^{n+1} f(x)dx$$

But since f is monotone decreasing over each interval $[n, n+1]$, the area under the graph of f is bounded above (resp. below) by the area under the constant function $x \mapsto f(n) = a_n$ (resp. $x \mapsto f(n+1) = a_{n+1}$). Thus we have, for every $n \geq 1$,

$$a_{n+1} \;\leq\; b_n \;\leq\; a_n.$$

Summing from $n = 1$ to ∞ and using the comparison test (see chapter 2), we get

$$S - a_1 \le T \le S.$$

Thus T converges as well.

$\square$

As a consequence we see that for any positive real number t, the series

$$S_t = \sum_{n=1}^{\infty} \frac{1}{n^t}$$

converges iff the improper integral

$$I_t = \int_1^{\infty} x^{-t} dx$$

converges. We have already seen that I_t is convergent iff $t > 1$. So the same holds for S_t. But recall that in the special case $t = 1$,we deduced the divergence of I_1 from that of S_1.

6 Differentiation, Properties, Tangents, Extrema

6.1 Derivatives

Let a be a real number and f a function defined on an interval around a. One says that f is **differentiable at** a iff the following limit exists:

$$L := \lim_{h\to 0} \frac{f(a+h) - f(a)}{h}.$$

When the limit exists, we will set

$$f'(a) \,=\, L.$$

We will also denote it sometimes by $\frac{df}{dx}(a)$.

Consider, for example, the case of a **linear function**

$$f(x) \,=\, mx + c,$$

whose graph is the **line of slope** m, passing through the point $(0, c)$. Since $f(x+h) = m(x+h) + c = f(x) + mh$, we have at any point a in $\mathbb{R}$,

$$\frac{f(x+h) - f(x)}{h} \,=\, \frac{mh}{h} \,=\, m,$$

which is independent of h. Hence it has the limit m as h approaches 0. Thus f is differentiable at $x = a$ with $f'(a)$ being the slope m.

In particular, when $m = 0$, the function f is just the **constant function** $x \mapsto c$, and the derivative is 0.

The next simple example to consider is the **quadratic function**

$$f(x) \,=\, \alpha x^2 + \beta x + \gamma,$$

with α, β, γ in $\mathbb{R}$. Then

$$\frac{f(x+h) - f(x)}{h} \,=\, \frac{(2\alpha x + \beta)h + \alpha h^2}{h} \,=\, 2\alpha + \beta + \alpha h,$$

which has the limit $2\alpha x + \beta$ as h tends to 0. Thus

$$\frac{df}{dx} \,=\, 2\alpha x + \beta.$$

This is L:eibnitz's notation, and it means that for any a,

$$f'(a) \,=\, \frac{df}{dx}(a) \,=\, 2\alpha a + \beta.$$

In particular, the **squaring function** $f(x) = x^2$ is differentiable everywhere with derivative $2x$. This is a superspecial case of the following important result on the **power function** x^t.

Proposition 6.1.1 *For any real number t, consider the function*

$$f(x) \;=\; x^t.$$

Then f is differentiable at any point a in $\mathbb{R}$, with

$$f'(a) \;=\; ta^{t-1}.$$

We will prove this in four stages. The first step is the following:

Proof when t is a positive integer m. Let $a, h \in \mathbb{R}$. By the binomial theorem,

$$(a+h)^m \;=\; \sum_{j=0}^{m} a^{m-j}h^j \;=\; a^m + ma^{m-1}h + \frac{m(m-1)}{2}a^{m-2}h^2 + \ldots h^m.$$

In other words,

$$\frac{(a+h)^m - a^m}{h} \;=\; ma^{m-1} + hg(a,h),$$

where

$$g(a,h) \;=\; \sum_{j=2}^{m} a^{m-j}h^{j-2} \;=\; \frac{m(m-1)}{2}a^{m-2} + \frac{m(m-1)(m-2)}{2}a^{m-3}h + \ldots + h^{m-2}.$$

Then $g(x,0) = \frac{m(m-1)}{2}a^{m-2}$, and so $hg(x,h)$ goes to 0 as h goes to 0. Consequently,

$$\lim_{h\to 0} \frac{(a+h)^m - a^m}{h} \;=\; ma^{m-1},$$

as asserted.

$\square$

Here is the next step.

Proof when $t = 1/n$ with n a positive integer. For any $a \in \mathbb{R}$, note that

$$h \;=\; z^n - w^n \quad \text{if} \quad z = (a+h)^{1/n},\ w = a^{1/n}.$$

Thus

$$\frac{(a+h)^{1/n} - a^{1/n}}{h} \;=\; \frac{z-w}{z^n - w^n}.$$

But we have the factorization

$$z^n - w^n \;=\; (z-w)(z^{n-1} + z^{n-2}w + z^{n-3}w^2 + \ldots + zw^{n-2} + w^{n-1}),$$

which can be verified by direct computation. Thus

$$\lim_{h\to 0} \frac{(a+h)^{1/n} - a^{1/n}}{h} = \lim_{h\to 0} \frac{z-w}{z^n - w^n} = \lim_{h\to 0} \frac{1}{z^{n-1} + z^{n-2}w + \ldots + zw^{n-2} + w^{n-1}}.$$

Since the function $x \mapsto x^{1/n}$ is continuous, the limit of $z^i w^j$ as h goes to 0 is simply w^{i+j}. We then obtain

$$\lim_{h\to 0} \frac{(a+h)^{1/n} - a^{1/n}}{h} = \frac{1}{nw^{n-1}} = \frac{1}{n} a^{\frac{1}{n}-1}.$$

which proves the assertion of Proposition 6.1.1 for $t = 1/n$.

□

We will come back to this proposition in the next section after deriving the *product rule*.

Proposition 6.1.2 *The functions $f(x) = \sin x$ and $g(x) = \cos x$ are differentiable at any point a in $\mathbb{R}$, and*

$$f'(a) = \cos a \quad \text{and} \quad g'(a) = -\sin a.$$

Proof. Recall the addition formula

$$\sin(x+y) = \sin x \cos y + \cos x \sin y.$$

Therefore

$$\frac{\sin(a+h) - \sin a}{h} = \sin a \frac{\cos h - 1}{h} + \cos a \frac{\sin h}{h}.$$

We have already seen that

$$\lim_{h\to 0} \frac{\sin h}{h} = 1 \quad \text{and} \quad \lim_{h\to 0} \frac{\cos h - 1}{h} = 0.$$

Consequently,

$$\lim_{h\to 0} \frac{\sin(a+h) - \sin a}{h} = \cos a,$$

as asserted.

Similarly, the addition formula for the cosine function, namely

$$\cos(x+y) = \cos x \cos y - \sin x \sin y,$$

implies

$$\frac{\cos(a+h) - \cos a}{h} = \cos a \frac{\cos h - 1}{h} - \sin a \frac{\sin h}{h}.$$

The expression on the right has the limit $-\sin a$ as h goes to 0. Done.

□

One can check that if f is differentiable at a, f must necessarily be continuous at a, but it is not sufficient. Indeed, the function $f(x) = |x|$ is continuous everywhere, but it is not integrable at $x = 0$. This is because for $h > 0$,

$$\frac{f(h) - f(0)}{h} = \frac{h - 0}{h} = 1,$$

while for $h < 0$,

$$\frac{f(h) - f(0)}{h} \;=\; \frac{-h - 0}{h} \;=\; -1.$$

So $\frac{f(h)-f(0)}{h}$ is 1, resp. -1, as h approaches 0 from the right, resp. left. So there is no unique limit and f is not differentiable at 0. Note, however, that f is differentiable everywhere else.

Let us end this section by looking at a couple of more examples. The first one below is **not differentiable at** $x = 0$:

$$f(x) \;=\; \begin{cases} x \sin\left(\frac{1}{x}\right), & \text{if} \quad x \neq 0 \\ 0, \quad \text{if} \quad x = 0 \end{cases}$$

Indeed, for any $h \neq 0$,

$$\frac{f(h) - f(0)}{h} \;=\; \frac{h \sin\left(\frac{1}{h}\right) - 0}{h} \;=\; \sin\left(\frac{1}{h}\right),$$

and $\sin\left(\frac{1}{h}\right)$ has no limit as h goes to 0; it fluctuates wildly between -1 and 1.

The second example is

$$g(x) \;=\; \begin{cases} x^2 \sin\left(\frac{1}{x}\right), & \text{if} \quad x \neq 0 \\ 0, \quad \text{if} \quad x = 0 \end{cases}$$

In this case

$$\frac{g(h) - g(0)}{h} \;=\; h \sin\left(\frac{1}{h}\right),$$

which approaches 0 as h goes to 0. So $g(x)$ is differentiable at $x = 0$ with

$$g'(0) \;=\; 0.$$

6.2 Rules of differentiation, consequences

The two basic results here are the following:

Theorem 6.2.1 *Let f, g be differentiable functions at some $a \in \mathbb{R}$. Then we have:*

*(i) (**Linearity**) For all α, β in $\mathbb{R}$, the function $\alpha f + \beta g$ is differentiable at a, with*

$$(\alpha f + \beta g)'(a) \;=\; \alpha f'(a) + \beta g'(a).$$

*(ii) (**Product rule**) The product function fg is differentiable at a, with*

$$(fg)'(a) \;=\; f'(a)g(a) + f(a)g'(a).$$

(iii) (**Quotient rule**) *If $g(x)$ is non-zero in an interval around a, then the ratio f/g is differentiable at a, with*

$$\left(\frac{f}{g}\right)(a) \;=\; \frac{f'(a)g(a) - f(a)g'(a)}{g(a)^2}.$$

Theorem 6.2.2 (**Chain rule**) *Let f be a differentiable function at some a in $\mathbb{R}$, g a differentiable function at $f(a)$. Then the composite function $g \circ f$ is differentiable at a, with*

$$(g \circ f)'(a) \;=\; g'(f(a)) \cdot f'(a).$$

Before giving proofs of these assertions, which will be done in the next section, let us note that as a **consequence**, **rational functions**, **trigonometric functions**, and **various combinations** of them are differentiable wherever they are defined.For example, the function

$$f(x) \;=\; \frac{x^{72} - 21x^3 + \sin^9(x^4 + 43x) - \cos^3(x^3 - 5)}{12\sin\sqrt{x} - 29x^4 - 4}$$

is differentiable at any x where the denominator is non-zero. A simpler, and more **commonly occurring example** is

$$g(x) \;=\; \tan x.$$

Since $\tan x = \sin x/\cos x$, we get by applying the *quotient rule* and Proposition 6.1.2,

$$\frac{d}{dx}\tan x \;=\; \frac{(\cos x)(\cos x) - (\sin x)(-\sin x)}{\cos^2 x} \;=\; \frac{1}{\cos^2 x} \;=\; \sec^2 x. \tag{6.2.3a}$$

Here we have used the fact that $\sin^2 x + \cos^2 x = 1$, and the formula is valid at any x where $\cos x$ is non-zero, i.e., at any real number x not equal to an odd integer multiple of $\pi/2$. Similarly we have

$$\frac{d}{dx}\cot x \;=\; -^2 x, \tag{6.2.3b}$$

$$\frac{d}{dx}\sec x \;=\; \sec x \tan x, \tag{6.2.3c}$$

and

$$\frac{d}{dx}x \;=\; -x\cot x, \tag{6.2.3d}$$

for all x where the function is defined.

Now we will complete the **proof of Proposition 6.1.1**.

In the previous section we proved the formula for the functions $f(x) = x^m$ and $g(x) = x^{1/n}$ for integers m, n with $m \geq 0$, $n > 0$. The next step is to look at the function

$$h(x) \;=\; x^{m/n},$$

for a rational number m/n, with $m, n > 0$. Then f is the composite of the functions f, g above, which are differentiable at every point in their respective domains. Thus for any $a > 0$, we have by the *chain rule*,

$$h'(a) \;=\; g'(f(a)) \cdot f'(a) \;=\; \left(\frac{1}{n}(f(a))^{1/n-1}\right) \cdot \left(ma^{m-1}\right),$$

which simplifies to

$$\frac{m}{n} a^{\frac{m}{n}-m} a^{+m-1} \;=\; \frac{m}{n} a^{\frac{m}{n}-1},$$

as desired. Now suppose

$$\phi(x) \;=\; x^{-m/n},$$

for some positive integers m, n. Then $\phi(x)$ is the reciprocal of the function $h(x) = x^{m/n}$ which we just looked at. By applying the *quotient rule*, we see that h is differentiable at any positive real number a and moreover,

$$\phi'(a) \;=\; \frac{-h'(a)}{h(a)^2} \;=\; -\frac{m}{n}\frac{a^{\frac{m}{n}-1}}{a^{\frac{2m}{n}}} \;=\; -\frac{m}{n} a^{-\frac{m}{n}-1},$$

as asserted.

The final step is to consider the function

$$f(x) \;=\; x^t,$$

for any *real* number t. But t is, by the construction of $\mathbb{R}$, the limit of a sequence $\{t_n\}$ of rational numbers, and for all $x \geq 0$ we have

(6.2.4)
$$x^t \;=\; \lim_{n\to\infty} x^{t_n}.$$

(This could be taken as the definition of x^t for real t.) We then have, for all $a > 0$,

(6.2.5)
$$\lim_{h\to 0} \frac{(a+h)^t - a^t}{h} \;=\; \lim_{h\to 0} \lim_{n\to\infty} \frac{(a+h)^{t_n} - a^{t_n}}{h}.$$

It can be shown that the order of taking the two limits can be reversed (see the later chapter of *infinite series of functions*). Thus the right hand side of (6.2.5) becomes

$$\lim_{n\to\infty} \lim_{h\to 0} \frac{(a+h)^{t_n} - a^{t_n}}{h} \;=\; \lim_{n\to\infty} \frac{d\{x^{t_n}\}}{dx}(a),$$

which simplifies to

(6.2.6)
$$\lim_{n\to\infty} t_n a^{t_n - 1} \;=\; ta^{t-1},$$

which is what we wanted to show.

□

6.3 Proofs of the rules

Now we will prove Proposition 6.2.1 and Theorem 6.2.2.

Proof of Proposition 6.2.1 By the linearity of limits (see Proposition 4.1.2) we have

$$\lim_{h\to 0} \frac{(\alpha f+\beta g)(a+h)-(\alpha f+\beta g)(a)}{h} \;=\; \alpha \lim_{h\to 0} \frac{f(a+h)-f(a)}{h} + \beta \lim_{h\to 0} \frac{g(a+h)-g(a)}{h},$$

which is

$$\alpha f'(a)+\beta g'(a),$$

proving (i).

Next,

$$\lim_{h\to 0} \frac{(fg)(a+h)-(fg)(a)}{h} \;=\; \lim_{h\to 0} \frac{(f(a+h)-f(a))g(a+h)+f(a)(g(a+h)-g(a))}{h},$$

which equals, by Proposition 4.1.2,

$$\lim_{h\to 0} \frac{(f(a+h)-f(a))g(a+h)}{h} + \lim_{h\to 0} \frac{f(a)(g(a+h)-g(a))}{h} \;=\; f'(a)g(a)+f(a)g'(a),$$

yielding (ii).

Now on to the *quotient rule*. We have

$$\lim_{h\to 0} \frac{\frac{f}{g}(a+h)-\frac{f}{g}(a))}{h} \;=\; \lim_{h\to 0} \frac{f(a+h)-f(a)}{hg(a+h)} - \frac{f(a)}{g(a)} \lim_{h\to 0} g(a)-g(a+h)hg(a+h),$$

which equals

$$\frac{f'(a)g(a)-f(a)g'(a)}{g(a)^2},$$

as asserted in (iii).

$\square$

Before beginning the proof of the *chain rule*, we need the following

Lemma 6.3.1 *Let f be a function defined in an interval around a real number a. Then the following are equivalent:*

(i) f is differentiable at a

(ii) There is a real number λ such that

$$\lim_{x\to a} \frac{f(x)-f(a)-\lambda(x-a)}{x-a} \;=\; 0.$$

When (ii) holds, the number λ is $f'(a)$.

Proof of Lemma. Note that f is differentiable at a iff the limit

$$L = \lim_{x \to a} \frac{f(x) - f(a)}{x - a}$$

exists, and when it exists, L is denoted by $f'(a)$. On the other hand,

$$\lim_{x \to a} \frac{f(x) - f(a) - \lambda(x - a)}{x - a} = \lim_{x \to a} \frac{f(x) - f(a)}{x - a} - \lambda = L - \lambda.$$

So L exists iff $L - \lambda$ exists and equals zero for some λ, namely for $\lambda = L$. It is now clear that (i) and (ii) are equivalent, and that when (ii) is satisfied, $\lambda = f'(a)$.

$\square$

Proof of Theorem 6.2.2. Let us put

$$b = f(a), \quad \phi = g \circ f,$$

$$F(x) = f(x) - f(a) - f'(a)(x - a), \; G(y) = g(y) - g(b) - g'(b)(y - b)$$

and

$$\Phi(x) = h(x) - h(a) - g'(b)f'(a)(x - a).$$

Then we have, by Lemma 6.3.1 above,

$$\lim_{x \to a} \frac{F(x)}{x - a} = 0 = \lim_{y \to b} \frac{G(y)}{y - b}. \tag{6.3.2}$$

In view of Lemma 6.3.1, it suffices to show:

$$\lim_{x \to a} \frac{\Phi(x)}{x - a} = 0. \tag{6.3.3}$$

But

$$\Phi(x) = g(f(x)) - g(b) - g'(b)\left(f'(a)(x - a)\right),$$

and since $f'(a)(x - a) = f(x) - f(a) - F(x)$, we get

$$\Phi(x) = [g(f(x)) - g(b) - g'(b)(f(x) - f(a))] + g'(b)F(x) = G(f(x)) + g'(b)F(x).$$

Thanks to (6.3.2), it then suffices to prove:

$$\lim_{x \to a} \frac{G(f(x))}{x - a} = 0. \tag{6.3.4}$$

We know that $\lim_{y \to b} \dfrac{|G(y)|}{|y - b|}$ is zero. So we can find, for every $\varepsilon > 0$, a $\delta > 0$ such that $|G(f(x))| < \varepsilon |f(x) - b|$ if $|f(x) - b| < \delta$. But since f is integrable at a, it is also continuous

there (see section 6.1). Hence $|f(x) - b| < \delta$ whenever $|x - a| < \delta_1$, for any small enough $\delta_1 > 0$. Hence

$$\begin{aligned} (f(x))| &< \varepsilon|f(x) - b| = \varepsilon|F(x) + f'(a)(x-a)| \\ &\leq \varepsilon(x)| + \varepsilon|f'(a)(x-a)|, \end{aligned}$$

by the triangle inequality. Since $\lim_{x \to a} \frac{F(x)}{x-a}$ is zero, we get

$$\lim_{x \to a} \frac{|G(f(x))|}{|x-a|} \leq \varepsilon \lim_{x \to a} \frac{|f'(a)(x-a)|}{|x-a|} = \varepsilon|f'(a)|.$$

Now (6.3.4) follows easily, and we are done.

6.4 Tangents

Recall from High school Geometry that the **equation of any non-vertical line L in the plane** is given by

(6.4.1) $$y = mx + c,$$

where m is the **slope** of the line and $(0, c)$ is a point on L. The equation of a vertical line is given by $x = c_1$ for some constant c_1, and this will not concern us here because the graph of any function $f(x)$ will never be a vertical line (check it!).

To know a (non-vertical) line L as above, it is sufficient to know the slope m and any point $P = (x_0, y_0)$ on L. (P need not be $(0, c)$.) Indeed, since the coordinates of P must satisfy the equation of L, we have

(6.4.2) $$y_0 = mx_0 + c,$$

and combining (6.4.1) and (6.4.2) we get

$$y - y_0 = m(x - x_0),$$

or equivalently,

(6.4.3) $$y = mx + (y_0 - mx_0).$$

So we can recover (6.4.1) by setting

$$c = y_0 - mx_0.$$

Now let f be a function defined in an interval I around a point a. Suppose f is differentiable at a. Then, as noted earlier, the slope of the line T_a which is tangent to the graph C of $y = f(x)$, $x \in I$, is given by

(6.4.4) $$m = f'(a).$$

Moreover, since the point $(a, f(a)$ belongs to T_a, the equation of T_a is, thanks to (6.4.3), given by

(6.4.5) $$y = f'(a)x + (f(a) - f'(a)a).$$

6.5 Extrema of differentiable functions

Recall that any **continuous function** f on a closed interval $[a, b]$ attains its extrema, i.e., there exist $c, d \in [a, b]$ such that

$$(6.5.1) \qquad f(c) \leq f(x) \leq f(d) \quad \forall\, x \in [a, b].$$

One says that c is a **minimum** and d a **maximum** of f on $[a, b]$.

While this is a very beautiful result, we do not see how one can go about finding the points c and d where the extremal values occur.

Things become nicer when f is **differentiable** on (a, b), and even nicer when f is **twice differentiable** on (a, b), i.e., when f is differentiable and also the derivative f' is differentiable on (a, b). Note that it does not make sense to require that f is differentiable at either of the end points, because to have differentiability at a point, the function needs to be defined in an open interval surrounding the point.

Definition 6.5.2 *A function f defined on a set S containing a real number c has a* **local minimum**, *resp.* **local maximum**, *at c iff we can find a $\delta > 0$ such that for all $x \in (c - \delta, c + \delta) \cap S$,*

$$f(x) \geq f(c) \quad (\text{resp. } f(x) \leq f(c).$$

Clearly, a local minimum need not be an **absolute minimum** of f on S, but it provides a candidate to be tested for being the absolute minimum (if one exists, as for a continuous f).

Here is a basic result.

Proposition 6.5.3 *Let f be a function differentiable at a real number c.*

(i) If c is a local minimum or a local maximum, then $f'(c) = 0$.

(ii) Suppose f is twice differentiable at c with $f'(c) = 0$ and $f''(c)0$ (resp. $f''(c) < 0$). Then c is a local minimum (resp. local maximum) for f.

Proof. *(i)* Suppose c is a local minimum. Then by definition, we have for all small enough h,

$$f(a + h) - f(a) \geq 0.$$

Thus the **right derivative** of f at c satisfies

$$L^+ = \lim_{h \to 0,\, h > 0} \frac{f(a + h) - f(a)}{h} \geq 0$$

and the **left derivative** of f at c satisfies

$$L^- = \lim_{h \to 0,\, h < 0} \frac{f(a + h) - f(a)}{h} \leq 0.$$

Since by assumption f is differentiable at c, L^+ and L^- both exist and equal $f'(c)$. This forces $f'(c)$ to vanish as claimed. The argument is similar when c is a local maximum. Hence (i).

For the next part we need the following

Lemma 6.5.4 *Let f be a differentiable function on an open interval I such that $f'(x)$ is > 0 (resp. < 0) for all x in I. Then f is a* **strictly increasing function**, *resp.* **strictly decreasing function**, *on I.*

Proof of Lemma. Assume $f'(x) > 0$ for all $x \in I$. Pick any two points $a, b \in I$ with $a < b$. We have to show that $f(a) < f(b)$. Since $f'(a) > 0$, $f(a+h) > f(a)$ for all small $h > 0$. If $f(a) \leq f(b)$, then $f(a+h) > f(b)$ for small $h > 0$, and this will force the graph to peak at some point between a and b, i.e., f will have a local maximum at some c in (a, b). By part (i) of this Proposition, $f'(c)$ will then be 0, contradicting the hypothesis that the derivative of f is positive on all of I. Hence $f(a) < f(b)$, and the assertion is proved. The argument is similar when $f'(x) < 0$ for all $x \in I$.

□

Proof of part (ii) of Proposition 6.5.3. Suppose $f'(c) = 0$ and $f''(c) > 0$. Then

$$f''(c) \ = \ \frac{f'(c+h)}{h} \ > \ 0.$$

Then $\frac{f(c+h)}{h}$ will need to be positive for all small h. When h is small and positive (resp. negative), this will mean that $f'(a+h)$ is positive (resp. negative). By the Lemma above, f will, for all small positive h, be strictly increasing in $(c, c+h)$ and strictly decreasing in $(c-h, c)$. So c will have to be a local maximum of f. By a similar argument, if $f'(c) = 0$ and $f''(c) < 0$, c will be a local maximum, as asserted.

□

Suppose f is a differentiable function around a point c. We will say that c is a **critical point** of f iff $f'(c) = 0$.

A critical point can be a local minimum or a local maximum or of neither type. For example, the function

$$f(x) \ = \ x^3$$

has $x = 0$ as the only critical point, as $f'(x) = 3x^2$. But 0 is neither a local maximum nor a local minimum.

6.6 The mean value theorem

Here is a very useful result to know:

The Mean Value Theorem *Let f be a continuous function on $[a, b]$. Suppose f is differentiable on (a, b). Then we can find a point c in (a, b) where the following holds:*

$$f'(c) \ = \ \frac{f(b) - f(a)}{b - a}.$$

A **special case** of this result, obtained by setting $f(a) = f(b)$, yields the following:

Rolle's Theorem *Let f be a continuous function on $[a, b]$. Suppose f is differentiable on (a, b). Then we can find a point c in (a, b) where $f'(c) = 0$.*

Prof of Rolle's theorem. Since f is continuous on $[a, b]$, we know that there exist c, d in $[a, b]$ for which (6.5.1) holds. If c or d lies in the open interval (a, b), then part (i) of Proposition 6.5.3 implies that the derivative vanishes there, proving Rolle's theorem. So we may assume that neither c nor d is in (a, b), i.e, $\{c, d\} = \{a, b\}$ as sets. Since by assumption, $f(a) = f(b)$, this forces us to have the maximum and the minimum value of f on $[a, b]$ to be $f(a)$ $(= f(b))$. The only possibility then is for f to be a constant function, forcing $f'(x)$ to be 0 at every $x \in (a, b)$.

$\square$

Proof of the Mean Value Theorem. Put

$$m = \frac{f(b) - f(a)}{b - a}, \tag{6.6.1}$$

which is the slope of the line joining $(a, f(a))$ and $(b, f(b))$. Put

$$g(x) = f(x) - m(x - a). \tag{6.6.2}$$

Then by additivity, g is continuous on $[a, b]$ and differentiable on (a, b). Moreover

$$g(a) = f(a),$$

and

$$g(b) = f(b) - m(b - a). \tag{6.6.3}$$

Combining (6.6.1) and (6.6.3), we get

$$g(b) = f(b) - (f(b) - f(a)) = f(a).$$

Hence

$$g(a) = g(b).$$

So we may apply Rolle's theorem and conclude that for some c in (a, b),

$$g'(c) = 0. \tag{6.6.4}$$

On the other hand, by the definition (6.6.2) of $g(x)$,

$$g'(c) = f'(c) - m, \tag{6.6.5}$$

The assertion of the Theorem now follows by combining (6.6.4), (6.6.5) and (6.6.1).

$\square$

We will end the section by noting certain facts concerning **convexity** and **concavity**.

Recall from the Homework assignment 3 that a point u lies in a closed interval $[c, d]$ iff we can write $u = (1-t)c + td$ for some real number t in $[0,1]$. Given a function f on $[a, b]$, we will say that f is **convex** on $[a, b]$ iff for every subinterval $[c, d]$ and for every t in $[0,1]$,, we have

$$f((1-t)c + td) \leq (1-t)f(c) + tf(d). \tag{6.6.6}$$

It is said to be **concave** if the inequality is reversed, i.e., iff

$$f((1-t)c + td) \geq (1-t)f(c) + tf(d). \tag{6.6.7}$$

Note that the function f is **linear** iff equality holds everywhere.

A useful result, which we will not prove, is the following:

Proposition 6.6.8 *Let f be a function which is differentiable on an open interval containing $[a, b]$. Then f is convex, resp. concave, on $[a, b]$ if f' is an increasing, resp. decreasing, function on this interval.*

Note that if f is twice differentiable, we can assert that f is convex, resp. concave, iff f'' is everywhere positive, resp. negative.

7 The Fundamental Theorems of Calculus, Methods of Integration

So far we have separately learnt the basics of integration and differentiation. But they are not unrelated. In fact, they are **inverse operations**. This is what we will try to explore in the first section, via the two fundamental theorems of Calculus. After that we will discuss the two main methods one uses for integrating somewhat complicated functions, namely **integration by substitution** and **integration by parts**.

7.1 The fundamental theorems

Suppose f is an integrable function on a closed interval $[a, b]$. Then we can consider the **signed area function A** on $[a, b]$ (relative to f) defined by the definite integral of f from a to x, i.e.,

$$A(x) = \int_a^x f(t)dt. \tag{7.1.0}$$

The reason for the **signed area** terminology is that f is not assumed to be ≥ 0, so apriori $A(x)$ could be negative.

It is extremely interesting to know how $A(x)$ varies with x. What conditions does one need to put on f to make sure that A is continuous, or even differentiable? The continuity part of the question is easy to answer.

Lemma 7.1.1 *Let f, A be as above. Then A is a continuous function on $[a, b]$.*

Proof. Let c be any point in $[a, b]$. Then f is continuous at c iff we have

$$\lim_{h\to 0} A(c+h) = A(c).$$

Of course, i taking the limit, we consider all small enough h for which $c+h$ lies in $[a, b]$, and then let h go to zero. By the *additivity* of the integral, we have (using (7.1.10)),

$$A(c+h) - A(c) = \int_{I(c,h)} f(t)dt,$$

where $I(c, h)$ denotes the closed interval between c and $c + h$. Clearly, $I(c, h)$ is $[c, c + h]$, resp. $[c + h, c]$, if h is positive, resp. negative. When h goes to zero, $I(c, h)$ shrinks to the point $\{c\}$, and so

$$\lim_{h\to 0} A(c+h) - A(c) = 0,$$

which is what we needed to show.

$\square$

The question of differentiability of A is more subtle, and the complete answer is given by the following important result:

Theorem 7.1.2 (**The first fundamental theorem of Calculus**) *Let f be an integrable function on $[a,b]$, and let A be the function defined by (7.1.0). Pick any point c in (a,b), and suppose that f is continuous at c. Then A is differentiable there and moreover,*

$$A'(c) \,=\, f(c).$$

Some would write this symbolically as

$$\frac{d}{dx}\int_a^x f(t)dt \,=\, f(x). \tag{7.1.3}$$

In plain words, this says that *differentiating the integral gives back the original* as long as the original function is continuous at the point in question.

Proof. To know if $A(x)$ is differentiable at c, we need to evaluate the limit

$$L \,=\, \lim_{h\to 0}\frac{A(c+h)-A(c)}{h}. \tag{7.1.4}$$

By the *additivity* of the definite integral, we have

$$A(c+h)-A(c) \,=\, \int_{I(c,h)} f(x)dx, \tag{7.1.5}$$

where $I(c,h)$ is as in the proof of Lemma 7.1.1.

Denote by $M(c,h)$, resp. $m(c,h)$, the *supremum*, resp. *infemum*, of the values of f over $I(c,h)$. Then the following bounds evidently hold:

$$hm(c,h) \,\le\, \int_{I(c,h)} f(x)dx \,\le\, hM(c,h). \tag{7.1.6}$$

Combining (7.1.4), (7.1.5) and (7.1.6), we get for all small h,

$$\lim_{h\to 0} m(c,h) \,\le\, L \,\le\, \lim_{h\to 0} M(c,h). \tag{7.1.7}$$

But by hypothesis, f is continuous at c. Then both $m(c,h)$ and $M(c,h)$ will tend to $f(c)$ as h goes to 0, which proves the Theorem in view of (7.1.7).

$\square$

Let f be any function on an open interval I. Suppose there is a differentiable function ϕ on I such that $\phi'(x) = f(x)$ for all x in I. Then we will call ϕ a **primitive** of f on I. Note that the primitive is not unique. Indeed, for any constant α, the function $\phi+\alpha$ will have the same derivative as ϕ. Intuitively, one feels immediately that the notion of a primitive

should be tied up with the notion of an integral. The following very important and oft-used result makes this expected relationship precise.

Theorem 7.1.8 (**The second fundamental theorem of Calculus**) *Suppose f, ϕ are functions on $[a,b]$, with f integrable on $[a,b]$ and ϕ a primitive of f on (a,b), with ϕ defined and continuous at the endpoints a, b. Then*

$$\phi(b) - \phi(a) \;=\; \int_a^b f(x)dx.$$

One can rewrite this, perhaps more expressively, as

$$\phi(b) - \phi(a) \;=\; \int_a^b \frac{d}{dx}\phi dx.$$

Proof. Choose any partition

$$P:\; a = t_0 < t_1 < \ldots < t_n = b,$$

and set, for each $j \in \{1, 2, \ldots, n\}$,

$$M_j \;=\; \sup(f([t_{j-1}, t_j])) \quad \text{and} \quad m_j \;=\; \inf(f([t_{j-1}, t_j])).$$

By definition,

$$\sum_{j=1}^{n} (t_j - t_{j-1})m_j \;\leq\; \int_a^b f(x)dx \;\leq\; \sum_{j=1}^{n} (t_j - t_{j-1})m_j. \tag{7.1.9}$$

On the other hand, the **Mean Value Theorem** gives us, for each j, a number c_j in $[t_{j-1}, t_j]$ such that

$$\phi'(c_j) \;=\; \frac{\phi(t_j) - \phi(t_{j-1})}{t_j - t_{j-1}}. \tag{7.1.10}$$

Since ϕ is by hypothesis the *primitive* of f on (a,b), $f(c_j) = \phi'(c_j)$ for each j. Moreover,

$$m_j \;\leq\; f(c_j) \;\leq\; M_j, \tag{7.1.11}$$

and

$$\sum_{j=1}^{n} \phi(t_j) - \phi(t_{j-1}) \;=\; \phi(b) - \phi(a). \tag{7.1.12}$$

Combining (7.1.10), (7.1.11) and (7.1.12), we obtain

$$\sum_{j=1}^{n} (t_j - t_{j-1})m_j \;\leq\; \phi(b) - \phi(a) \;\leq\; \sum_{j=1}^{n} (t_j - t_{j-1})m_j. \tag{7.1.13}$$

Since (7.1.9) and (7.1.13) hold for *every* partition P, and since f is integrable on $[a,b]$, the assertion of the Theorem follows.

□

7.2 The indefinite integral

Suppose ϕ is a primitive of a function f on an open interval I. It is not unusual to set, following Leibniz,

$$\int f(x)dx = \phi(x). \tag{7.2.1}$$

This is called an **indefinite integral** because there are no limits and ϕ is non-unique. So one can think of such an indefinite integral as a function of x which is unique only up to addition of an arbitrary constant. One has, in other words, an equality for all scalars C

$$\int f(x)dx = \int f(x)dx + C.$$

It could be a bit unsettling to work with such an indefinite, nebulous function at first, but one learns soon enough that it is a useful concept to be aware of.

In many Calculus texts one finds formulas like

$$\int \cos x dx = \sin x + C$$

and

$$\int \frac{1}{x} dx = \log x + C.$$

All they mean is that $\sin x$ and $\log x$ are the primitives of $\cos x$ and $\frac{1}{x}$, i.e.,

$$\frac{d}{dx} \sin x = \cos x$$

and

$$\frac{d}{dx} \log x = \frac{1}{x}.$$

Of course the situation is completely different in the case of definite integrals.

7.3 Integration by substitution

There are a host of techniques which are useful in evaluating various definite integrals. We will single out two of them in this chapter and analyze them. The first one is the method of substitution, which one should always try first before trying others.

Theorem 7.3.1 *Let $[a, b]$ be a closed interval and g a function differentiable on an open interval containing $[a, b]$, with g' continuous on $[a, b]$. Also let f be a continuous function on $g([a, b])$. Then we have the identity*

$$\int_a^b f(g(x))g'(x)dx = \int_{g(a)}^{g(b)} f(u)du.$$

Proof. Let ϕ denote a primitive of f, which exists because the continuity assumption on f makes it integrable on $[a, b]$. Then we have, by the second fundamental theorem of Calculus,

$$\int_{g(a)}^{g(b)} f(u)du = \phi(g(b)) - \phi(g(a)) = (\phi \circ g)(b) - (\phi \circ g)(a). \tag{7.3.2}$$

On the other hand, by the *chain rule* applied to the composite function $\phi \circ g$, we have

$$(\phi \circ g)'(x) = (\phi' \circ g)(x) \cdot g'(x) = (f \circ g)(x) \cdot g'(x).$$

Consequently,

$$\int_a^b f(g(x))g'(x)dx = \int_a^b (\phi \circ g)'(x)dx. \tag{7.3.3}$$

Applying the second fundamental theorem of Calculus again, the right hand side of (7.3.3) is the same as

$$(\phi \circ g)(b) - (\phi \circ g)(a). \tag{7.3.4}$$

The Theorem now follows by combining (7.3.2), (7.3.3) and (7.3.4).

$\square$

Before giving some examples let us note that powers of $\sin x$ and $\cos x$, as well as polynomials, are differentiable on $\mathbb{R}$ with continuous derivatives. In fact we can differentiate them any number of times; one says they are *infinitely differentiable*. The same holds for ratios of such functions or their combinations, as long as the denominator is non-zero in the interval of interest.

Examples 7.3.5: (**1**) Let

$$I = \int_0^{\pi/2} sin^3 x \cos x dx.$$

Thanks to the remark above on the *infinite differentiability* of the functions in the *integrand*, we are allowed to apply Theorem 7.3.1 here, with

$$g(x) = \sin x \quad \text{and} \quad f(u) = u^3.$$

Then, since $g'(x) = \cos x$ (as proved earlier, $g(0) = 0$ and $g(\pi/2) = 1$, we obtain

$$I = \int_0^1 u^3 du = \frac{1}{4}.$$

(**2**) Put

$$I = \int_0^{\pi/4} \cos^2 x dx.$$

Recall that

$$\cos^2 x - \sin^2 x = \cos 2x.$$

Since $\cos^2 x + \sin^2 x = 1$, we get

$$\cos^2 x = \frac{1 + \cos 2x}{2}.$$

Using this and the easy integral $\int_0^{\pi/4} dx = \pi/4$, we get

$$I = \frac{\pi}{8} + J, \quad \text{with} \quad J = \frac{1}{2} \int \cos 2x dx.$$

Put $g(x) = 2x$ and $f(u) = \cos u$, which are both infinitely differentiable on all of $\mathbb{R}$, and use Theorem 7.3.1. to conclude that, since $g'(x) = 2$, $g(0) = 0$ and $g(\pi/4) = \pi/2$,

$$J = \frac{1}{4} \int_0^{\pi/4} \cos u du = \frac{\sin(\pi/4) - \sin 0)}{4} = \frac{1}{4\sqrt{2}},$$

This implies that

$$I = \frac{\pi}{8} + \frac{1}{4\sqrt{2}} = \frac{\pi + \sqrt{2}}{8}.$$

(**3**) Evaluate

$$I = \int_0^1 \sqrt{1 - x^2} dx.$$

Here we use the substitution theorem in the reverse direction. The basic idea is that $\sqrt{1 - x^2}$ would simplify if x were $\sin t$ or $\cos t$. Put

$$g(t) = \sin t \quad \text{and} \quad f(u) = \sqrt{u}.$$

Then g is differentiable everywhere with $g'(t) = \cos t$ being continuous on $[0, 1]$. We chose the interval $[0, 1]$ because $g(0) = 0$ and $g(\pi/2) = 1$, giving us the limits of integration of I. Also, f is continuous on $g([0, \pi/2]) = [0, 1]$. (At the end point 0, the continuity of f means it is right continuous there. This is good, because f is not defined to the left of 0.) So we have satisfied all the hypotheses of Theorem 7.3.1 and we may apply it to get

$$I = \int_0^{\pi/2} f(g(t))g'(t)dt = \int_0^{\pi/2} \sqrt{1 - \sin^2 t} \cos t dt.$$

But

$$\sqrt{1-\sin^2 t} \;=\; \sqrt{\cos^2 t} \;=\; |\cos t|,$$

which is just $\cos t$, because the cosine function is non-negative in the interval $[0, \pi/2]$. Hence

$$I \;=\; \int_0^{\pi/2} \cos^2 t dt.$$

We just evaluated the integral of $\cos^2 t$ in the previous example, albeit with different limits. In any case, proceeding as in that example, we get

$$I \;=\; \frac{\pi}{4} + \frac{\sin(\pi/2) - \sin 0}{4} \;=\; \frac{\pi - 1}{4}.$$

7.4 Integration by parts

Some consider this the most important theorem of Calculus. Its use is pervasive.

Theorem 7.4.1 *Let $[a, b]$ be a closed interval and let f, g be differentiable functions in an open interval around $[a, b]$ such that f', g' continuous on $[a, b]$. Then we have*

$$\int_a^b f(x)g'(x)dx \;=\; f(x)g(x)|_a^b - \int_a^b f'(x)g(x)dx.$$

Here $f(x)g(x)|_a^b$ denotes $f(b)g(b) - f(a)g(a)$.

Proof. By the *product rule*,

$$(fg)'(x) \;=\; f(x)g'(x) + f'(x)g'(x) \tag{7.4.2}$$

for all x where f and g are both differentiable. Subtracting $f'(x)g(x)$ from both sides and integrating over $[a, b]$ we get

$$\int_a^b f(x)g'(x)dx \;=\; \int_a^b (fg)'(x)dx - \int_a^b f'(x)g(x)dx. \tag{7.4.3}$$

But by the second fundamental theorem of Calculus,

$$\int_a^b (fg)'(x)dx \;=\; (fg)(b) - (fg)(a). \tag{7.4.4}$$

The assertion now follows by combining (7.4.3) and (7.4.4).

□

Example 7.4.5: Evaluate, for *any* integer $n \geq 0$,

$$I_n = \int_0^{\pi/2} \cos^n x dx.$$

First note that

$$I_0 = \int_0^{\pi/2} 1 \cdot dx = \frac{\pi}{2}$$

and

$$I_1 = \int_0^{\pi/2} \cos x dx = \sin(\pi/2) - \sin 0 = 1,$$

because $\sin(\pi/2) = 1$ and $\sin 0 = 0$.

We have already solved the $n = 2$ case in the previous section using substitution, but we will not use it here. Integration by parts is more powerful!

So we may suppose that $n > 1$. Put

$$f(x) = \cos^{n-1} x \quad \text{and} \quad g(x) = \sin x.$$

Then, as noted earlier, f and g are infinitely differentiable on all of $\mathbb{R}$, with

$$f'(x) = (n-1)\cos^{n-2} x \cdot (-\sin x) \quad \text{and} \quad g'(x) = \cos x,$$

where the first formula comes from the chain rule. Now we may apply Theorem 7.4.1 and obtain

$$I_n = \left(\cos^{n-1}(x) \sin x\right)|_0^{\pi/2} - (n-1) \int_0^{\pi/2} cos^{n-2} x(-\sin x) \cdot \sin x dx.$$

Note that $n - 1 \neq 0$ as $n > 1$ and $\cos\theta \sin\theta$ is 0 if θ is 0 or $\pi/2$. Therefore the first term on the right is zero, and we get

$$I_n = (n-1) \int_0^{\pi/2} \cos^{n-2} x \sin^2 x dx.$$

Since $\sin^2 x = 1 - \cos^2 x$, we get

$$I_n = (n-1)I_{n-2} + (n-1)I_n,$$

which translates into the neat **recursive relation**

$$I_n = \frac{n-1}{n} I_{n-2}.$$

In particular,

$$I_2 = \frac{1}{2}I_0 = \frac{\pi}{4},\ I_4 = \frac{3}{4}I_2 = \frac{3\pi}{16}, \ldots$$

and

$$I_3 = \frac{2}{3}I_1 = \frac{2}{3},\ I_5 = \frac{4}{5}I_3 = \frac{8}{15}.$$

It will be left as a nice exercise for the reader to find exact formulae for I_{2n} and I_{2n-1}.

8 Factorization of polynomials, Integration by partial fractions

We know by now how to integrate any polynomial. What about rational functions? This is much harder. To approach it systematically, we first need some basic facts about polynomials and their factorizations. This leads to what is called a **partial fraction expansion** of the reciprocal $\frac{1}{f(x)}$ of a polynomial $f(x)$. Using this we can reduce the study of the general case to the following three cases:

$$\int \frac{dx}{(x-a)^m} \tag{8.0.1}$$

$$\int \frac{dx}{(x^2+bx+c)^m} \tag{8.0.2}$$

and

$$\int \frac{xdx}{(x^2+bx+c)^m}. \tag{8.0.3}$$

Here a, b, c are arbitrary scalars and m any positive integer.

It turns out that we can compute (8.0.1) and (8.0.3) without trouble (see section 8.4) by using *substitution* and the logarithm. To evaluate (8.0.2) even for $m = 1$, however, will require that we understand the *arctangent* function, which we will be do in chapter 9. One tackles the $m > 1$ case of (8.1.2) by a reduction process, as in the evaluation of the integral of $\cos^n x$.

8.1 Long division, roots

Suppose F is any field, for example $\mathbb{R}$, $\mathbb{C}$ or $\mathbb{Q}$. By a **polynomial of degree** n **with coefficients in** F, we mean a function of the form

$$f(x) \;=\; a_0 + a_1x + a_2x^2 + \ldots + a_nx^n \quad \text{with} \quad a_j \in F\,(\forall j) \quad \text{and} \quad a_n \neq 0. \tag{8.1.1}$$

One writes $\deg(f)$ for the degree n of f. Note that if two g is a polynomial of degree m with coefficients in F, then

$$\deg(fg) \;=\; \deg(f) + \deg(g), \tag{8.1.2}$$

while

$$\deg(f+g) \;\leq\; \max\{\deg(f), \deg(g)\}.$$

One defines the **degree of the zero polynomial** to be $-\infty$. This is forced upon us so as to have (8.1.2) when $f = 0$ and g arbitrary. (Think about it!)

Everyone is used to doing **long division** of polynomials in High School, and what results is the following useful result, called the **Euclidean algorithm for polynomials**:

Theorem 8.1.3 *Let f, g be polynomials with coefficients in F. Then we can find other polynomials Q, called the* **quotient polynomial**, *and R, called the* **remainder**, *such that*

$$f(x) = Q(x)g(x) + R(x) \quad \text{with} \quad \deg(R) < \deg(g).$$

We will say that g **divides** f, or that g **is a factor of** f, when there is no remainder, i.e., when R is the zero polynomial, in which case we have the **factorization**

$$f(x) = Q(x)g(x).$$

A number α in F will be called a **root** of a polynomial f iff we have $f(\alpha) = 0$. The following useful Lemma tells us the relationship between having a root and being factorizable.

Lemma 8.1.4 *Given a scalar α in F and a non-constant polynomial f with coefficients in F, the folowing are equivalent:*

(i) *α is a root of f;*

(ii) *$x - \alpha$ divides $f(x)$.*

Proof. Applying Theorem 8.1.3 with $g(x) = x - \alpha$, we get

$$f(x) = Q(x)(x - \alpha) + R(x) \tag{8.1.5}$$

with $\deg(R)$ <$\deg(g) = 1$. Then the only possibility is for R to be a **constant polynomial**, i.e., a scalar c. Then, plugging in $x = \alpha$ in (8.1.5), we obtain

$$f(\alpha) = Q(\alpha)(\alpha - \alpha) + c = c,$$

which implies that $f(\alpha)$ is zero iff c is zero, which happens iff $x - \alpha$ divides f.

$\square$

To paraphrase this Lemma, **every time we find a root we can factor!**

A polynomial f will be called **irreducible over** F iff f has **no factor** g with

$$0 < \deg(g) < \deg(f). \tag{8.1.5}$$

Lemma 8.1.6

(i) *Any* **linear polynomial** *f, i.e., one of degree 1, is irreducible.*

(ii) *The polynomial $f(x) = ax^2 + bx + c$ is irreducible over $\mathbb{R}$ iff the* **discriminant** *$D = b^2 - 4ac$ is negative.*

The most important example of (ii) is of course $x^2 + 1$.

Proof. Part (i) is clear from the condition (8.1.5). So let us prove part (ii). We know from High School mathematics that the roots of the quadratic polynomial $f(x) = ax^2+bx+c$ are given by

$$\alpha_{\pm} \,=\, -\frac{-b \pm \sqrt{D}}{2a},$$

with $D = b^2 - 4ac$. These roots always exist in $\mathbb{C}$. But for them to exist in $\mathbb{R}$, it is necessary and sufficient that D have a square-root in $\mathbb{R}$, i.e., that D be non-negative in $\mathbb{R}$. (The roots are not distinct iff $D = 0$.)

8.2 Factorization over $\mathbb{C}$

The most important result over $\mathbb{C}$, which is the reason people are so interested in working with complex numbers, is the following:

Theorem 8.2.1 (**The Fundamental Theorem of Algebra**) *Every non-constant polynomial with coefficients in $\mathbb{C}$ admits a root in $\mathbb{C}$.*

We will not prove this result here. But one should become aware of its existence if it is not already the case! We will now give an important consequence.

Corollary 8.2.2 Let f be a polynomial of degree $n \geq 1$ with $\mathbb{C}$-coefficients. Then there exist complex numbers $\alpha_1, \ldots, \alpha_r$, with $\alpha_i \neq \alpha_j$ if $i \neq j$, positive integers $m_1, \ldots, m_r$, and a scalar c, such that

$$f(x) \,=\, c\prod_{j=1}^{r}(x - \alpha_j),$$

and

$$\sum_{j=1}^{r} m_j \,=\, n.$$

In other words, any non-constant polynomial f with $\mathbb{C}$-coefficients **factorizes completely into a product of linear factors**. For each $j \leq r$, the associated positive integer m_j is called the **multiplicity** of α_j as a root of f, which means concretely that m_j is the highest power of $(x - \alpha_j)$ dividing $f(x)$.

Proof of Corollary. Let $n \geq 1$ be the degree of f and let a_n be the non-zero **leading coefficient**, i.e, the coefficient of x^n (see (8.1.1)). Let us set

$$c \,=\, a_n. \tag{8.2.3}$$

If $n = 1$,

$$f(x) \,=\, a_1x + a_0 \,=\, c(x - \alpha_1) \quad \text{with} \quad \alpha_1 \,=\, \frac{a_0}{a_1}.$$

So we are done in this case by taking $r = 1$ and $m_1 = 1$.

Now let $n > 1$ and assume by induction that we have proved the assertion for all $m < n$, in particular for $m = n - 1$. By Theorem 8.2.1, we can find a root, call it β, of f. By applying (8.1.4), we may then write

$$f(x) \,=\, (x - \beta)h(x), \tag{8.2.4}$$

for some polynomial $h(x)$ necessarily of degree $n - 1$. The leading coefficients of f are evidently the same. By induction we may write

$$h(x) \,=\, c\prod_{i=1}^{s}(x - \alpha_i)^{k_i},$$

for some roots $\alpha_1, \ldots, \alpha_s$ of h with respective multiplicities $n_1, \ldots, n_s$, so that

$$\sum_{i=1}^{s} k_i \,=\, n - 1.$$

But by (8.2.4), every root of h is also a root of f, and the assertion of the Corollary follows. □

8.3 Factorization over $\mathbb{R}$

The best way to understand polynomials f with real coefficients is to first look at their complex roots and then determine which ones of them could be real. To this end recall first the baby fact that a complex number $z = u + iv$ is real iff z equals its **complex conjugate** $\overline{z} = u - iv$, where $i = \sqrt{-1}$.

Proposition 8.3.1 *Let*

$$f(x) \,=\, a_0 + a_1x + \ldots + a_nx^n \quad \text{with} \quad a_j \in \mathbb{R}\,\forall j \le n \quad \text{and} \quad a_n \neq 0,$$

for some $n \ge 1$. Suppose α is a **complex root** *of f. Then $\overline{\alpha}$ is also a root of f. In particular, if r denotes the number of real roots of f and s the non-real (complex) roots of f, then we must have*

$$n \,=\, r + 2s.$$

We get the following consequence, which we proved earlier using the *Intermediate value theorem.*

Corollary 8.3.2 *Let f be a real polynomial of odd degree. Then f must have a real root.*

Proof of Proposition. Let α be a complex root of f. Recall that for all complex numbers z, w,

$$\overline{zw} \,=\, \overline{zw} \quad \text{and} \quad \overline{z + w} \,=\, \overline{z} + \overline{w}. \tag{8.3.3}$$

Hence for any $j \leq n$,

$$(\overline{\alpha})^j = \overline{\alpha^j}.$$

Moreover, since $a_j \in \mathbb{R}$ $(\forall j)$, $\overline{a}_j = a_j$, and therefore

$$a_j(\overline{\alpha})^j = \overline{a_j \alpha^j}.$$

Consequently, using (8.3.3) again, we get

$$f(\overline{\alpha}) = \sum_{j=0}^{n} a_j(\overline{\alpha})^j = \overline{f(\alpha)}. \tag{8.3.4}$$

But α is a root of f (which we have not used so far), $f(\alpha)$ vanishes, as does its complex conjugate $\overline{f(\alpha)}$. So by (8.3.4), $f(\overline{\alpha})$ is zero, showing that $\overline{\alpha}$ is a root of f.

So the **non-real roots** come in **conjugate pairs**, and this shows that n minus the number r, say, of the **real roots** is even. Done.

□

Given any complex number z, we have

$$z + \overline{z},\ z\overline{z} \in \mathbb{R}. \tag{8.3.5}$$

This is clear because both the **norm** $z\overline{z}$ and the **trace** $z + \overline{z}$ are unchanged under complex conjugation.

Proposition 8.3.6 *Let f be a real polynomial of degree $n \geq 1$ with real roots $\alpha_1, \ldots, \alpha_k$ with multiplicities $n_1, \ldots, n_k$, and non-real roots $\beta_1, \overline{\beta}_1, \ldots, \beta_\ell, \overline{\beta}_\ell$ with multiplicities $m_1, \ldots, m_\ell$ in $\mathbb{C}$. Then we have the factorization*

$$f(x) = c\prod_{i=1}^{k}(x - \alpha_i)^{n_i} \cdot \prod_{j=1}^{\ell}(x^2 + b_j x + c_j)^{m_j}, \tag{$*$}$$

where for each $j \leq \ell$,

$$b_j = -(\beta_j + \overline{\beta}_j) \quad \text{and} \quad c_j = \beta_j\overline{\beta}_j,$$

Each of the factors occurring in $(*)$ is a real polynomial, and the polynomials $x - \alpha_i$ and $x^2 + b_j x + c_j$ are all irreducible over $\mathbb{R}$.

Proof. In view of Corollary 8.2.2 and Proposition 8.3.1, the only thing we need to prove is that for each $j \leq \ell$, the polynomial

$$h_j(x) = x^2 b_j + c_j$$

is real and irreducible over $\mathbb{R}$. The reality of the coefficients $b_j = -(\beta_j + \overline{\beta}_j)$ and $c_j = \beta_j\overline{\beta}_j$ follows from (8.3.5). Suppose it is reducible over $\mathbb{R}$. Then we can write

$$h_j(x) = (x - t_j)(x - t'_j)$$

for some real numbers t_j, t'_j. On the other hand $\beta_j, \overline{\beta}_j$ are roots of h_j. This forces the equality of the sets $\{t_j, t'_j\}$ and $\{\beta_j, \overline{\beta}_j\}$, contradicting the fact that β_j is non-real. So h_j must be irreducible over $\mathbb{R}$.

□

8.4 The partial fraction decomposition

Here is the main result.

Theorem 8.4.1 *Let*

$$g(x) \;=\; \prod_{i=1}^{k}(x-\alpha_i)^{n_i}\cdot\prod_{j=1}^{\ell}(x^2+b_jx+c_j)^{m_j},$$

where the α_i, b_j, c_j *are real, and the* n_i, m_j *are positive integers. Then there exist real numbers* $A_i^{(p)}, B_j^{(q)}, C_j^{(q)}$, *with* $1\le i\le k$, $1\le p\le n_i$, $1\le j\le \ell$ *and* $1\le q\le m_j$, *such that*

$$\frac{1}{g(x)} \;=\; \sum_{i=1}^{k}\sum_{p=1}^{n_i}\frac{A_j^{(p)}}{(x-\alpha_i)^{(p)}} \;+\; \sum_{j=1}^{\ell}\sum_{q=1}^{m_j}\frac{B_j^{(q)}x+C_j^{(q)}}{(x^2+b_jx+c_j)^{(q)}}. \tag{8.4.2}$$

We will not prove this here. But here is the basic idea of the proof. Cross multiply (8.4.2) and get a polynomial equation of degree $n=\sum_{i=1}^{k}n_i+\sum_{j=1}^{\ell}m_j$ (which is the degree of g) where the coefficients involve the n indeterminates $A_i^{(p)}, B_j^{(q)}, C_j^{(q)}$. One solves for them by comparing the coefficients of x^i, for $1\le i\le n$. This results in an $n\times n$ **linear system**, i.e., a system of n linear equations in the n unknowns. In Ma1b you will learn to determine when such a linear system has a solution.

Let us try to understand this procedure in the simple case when

$$g(x) \;=\; (x-\alpha)^2(x^2+bx+c).$$

We want to show that there exist numbers A^1, A^2, B, C such that

$$\frac{1}{g(x)} \;=\; \frac{A^{(1)}}{x-\alpha}+\frac{A^{(2)}}{(x-\alpha)^2}+\frac{Bx+C}{x^2+bx+c}.$$

Cross multiplying, this gives the equation

$$1 \;=\; A^{(1)}(x-\alpha)(x^2+bx+c)+A^{(2)}(x^2+bx+c)+(Bx+C)(x-\alpha)^2.$$

Multiplying the right hand side out, we obtain

$$1 \;=\; A^{(1)}(x^3+(b-\alpha)x^2+(c-\alpha)x-c\alpha)+A^{(2)}(x^2+bx+c)+B(x^3-2\alpha x^2+\alpha^2x)+C(x^2-2\alpha x+\alpha^2).$$

Comparing coefficients, we get

$$(i)\qquad A^{(1)}+B \;=\; 0,\; A^{(1)}(b-\alpha)+A^{(2)}-2B\alpha+C \;=\; 0,$$

$$(ii)\qquad A^{(1)}(c-\alpha)+A^{(2)}b+B\alpha^2-2C\alpha \;=\; 0,\quad \text{and } -A^{(1)}c\alpha+A^{(2)}c+C\alpha^2 \;=\; 1.$$

This gives four linear equations in four unknowns, namely in $A^{(1)}, A^{(2)}, B$ and C. The equations (i) imply

$$(iii) \qquad A^{(1)} = -B \quad \text{and} \quad A^{(2)} = -A^{(1)}(b-\alpha) + 2B\alpha - C = B(b+\alpha) - C.$$

Eliminating $A^{(1)}, A^{(2)}$ from (ii) using (iii), we get

$$(iv) \qquad B(\alpha^2 + (b+1)\alpha + b^2 - c) - C(b + 2\alpha) = 0$$

and

$$(v) \qquad B(b+2\alpha)c + C(\alpha^2 - c) = 1.$$

It can be checked that the linear equations (iv), (v) are independent, so that we can solve for B, C in terms of α, b, c. Then we can find $A^{(i)}, i = 1, 2$ by using (iii).

To have a **numerical example**, take

$$\alpha = 1, b = 0, c = 1.$$

Then (iv) becomes $B - 2C = 0$ and (v) becomes $2B = 1$, so the solution we seek is given by

$$B = \frac{1}{2},\ C = \frac{1}{4},\ A^{(1)} = -\frac{1}{2},\ A^{(2)} = \frac{1}{2}.$$

Therefore

$$\frac{1}{(x-1)^2(x^2+1)} = -\frac{1}{2(x-1)} + \frac{1}{2(x-1)^2} + \frac{2x+1}{4(x^2+1)}.$$

8.5 Integration of rational functions

Let us begin discussing a simple situation. The numerical example at the end of section 8.4 implies, by the additivity of the integral, that

$$(8.5.1) \qquad \int \frac{dx}{(x-1)^2(x^2+1)} = I_1 + I_2 + I_3,$$

where

$$I_1 = -\frac{1}{2}\int \frac{dx}{x-1},$$

$$I_2 = \frac{1}{2}\int \frac{dx}{(x-1)^2}$$

and

$$I_3 = \frac{1}{4}\int \frac{2x+1}{x^2+1}dx.$$

We will see in the next chapter that

$$I_1 = -\frac{1}{2}\log|x-1| + C.$$

Using substitution and the knowledge of the integral of x^t, we get

$$I_2 = -\frac{1}{2(x-1)} + C.$$

And

$$I_3 = I_{3,1} + I_{3,2},$$

where

$$I_{3,1} = \frac{1}{4}\int \frac{2x}{x^2+1}dx = \frac{1}{4}\log(x^2+1) + C,$$

which was evaluated by using the substitution $u = x^2 + 1$, and

$$I_{3,2} = \frac{1}{4}\int \frac{dx}{x^2+1},$$

which we will be able to evaluate in the next chapter when we study the arctan function.

Suppose we want to integrate a **general rational function**. We have the following result.

Proposition 8.5.2 *Let $\frac{f(x)}{g(x)}$ be a rational function, i.e., a quotient of polynomials $f(x), g(x)$, with real coefficients. Then the (indefinite)*

$$I = \int \frac{f(x)}{g(x)}dx$$

can be written as a real linear combination of integrals of the form (8,0,1), (8,0,2) and (8,0,3) and the integral of a polynomial (which shows up only when deg(f) $\geq$deg(g)).

Proof. Thanks to Proposition 8.3.6 and Theorem 8.4.1, we can write I as a linear combination of integrals of the form

$$I_1 = \int \frac{h(x)}{(x-a)^m}dx \tag{8.5.3}$$

and

$$I_2 = \int \frac{h(x)}{(x^2+bx+c)^m}dx,$$

where $h(x)$ denotes a polynomial with real coefficients. In fact, $h(x)$ is in I_1, resp. I_2, a multiple of $f(x)$, resp. $(Ax+B)f(x)$ for some A, B.

There is nothing to prove if $f(x)$ is a constant. So let us take the degree of f to be ≥ 1. The Proposition is clearly a consequence of the following

Lemma 8.5.4 *Let $\phi(x)$ be a real polynomial of degree ≥ 1, and let a, b, c be real numbers with $a, b \neq 0$. Then*

(i) We can write $\phi(x)$ as a polynomial in $(x-a)$ with real coefficients.

(ii) If $\phi(x)$ has degree ≥ 2, then we can write

$$\phi(x) = \sum_{j=0}^{r} \lambda_j(x)(bx+c)^j,$$

where each $\lambda_j(x)$ is a real polynomial of degree ≤ 1.

Proof of Lemma 8.5.4. Let the degree of $\phi(x)$ be n.
(i) The assertion is obvious if $n = 1$, So take n to be > 1 and assume by induction that the assertion holds for $n-1$. Using Theorem 8.1.3, we can write

$$\phi(x) = Q(x)(x-a) + c_1, \tag{8.5.5}$$

where c_1 is a constant. Since $Q(x)$ is of degree $n-1$, we may apply the inductive hypothesis and conclude that $Q(x)$ is a real polynomial in $x-a$. Then (8.5.5) shows that $\phi(x)$ is also a polynomial in $x-a$, as claimed. Done.
(ii) We will apply the *Principle of Induction* to the set of all integers ≥ 2. Suppose $n = 2$, with $\phi(x) = Ax^2+Bx+C$, $A \neq 0$. Then $\phi(x)$ can be written as $A(x^2+bx+c)+((B-Ab)x+(C-Ac))$, so the assertion holds with $\lambda_0(x) = (B-Ab)x+(C-Ac)$ and $\lambda_1(x) = A$. So take n to be greater than 2 and assume by induction that the assertion holds for all $m < n$. Now applying Theorem 8.1.3 again, we may write

$$\phi(x) = Q(x)(x^2+bx+c) + \lambda_0(x), \tag{8.5.6}$$

where $\lambda_0(x)$ is a real polynomial of degree < 2. Since the degree of $Q(x)$ is of degree $n-2$, we may apply the inductive hypothesis and conclude that

$$Q(x) = \sum_{i=0}^{k} \mu_i(x)(x^2+bx+c),$$

with each $\mu_i(x)$ if a real polynomial of degree ≤ 1. Combining this with (8.5.6), we get what we want with $r = k+1$ and $\lambda_j(x) = \mu_{j-1}(x)$ for each $j \geq 1$.

$\square$

The next key step is to **complete the square**. Explicitly, we can, given any pair of real numbers b, c, write

$$x^2+bx+c = (x+\frac{b}{2})^2 + (c-\frac{b^2}{4}). \tag{8.5.7}$$

When x^2+bx+c is irreducible over $\mathbb{R}$, which we may assume to be the case, thanks to Proposition 8.3.6, its discriminant b^2-4c is necessarily negative, and so $c-\frac{b^2}{4}$ is positive and thus can be expressed as e^2, for a positive real number e. Consequently, by using the substitution $u = x+\frac{b}{2}$, which gives $u'(x) = 1$, we can transform the integrals (8.0.2) and (8.0.3) into linear combinations of integrals of the following form:

$$I_1 = \int \frac{du}{(u^2+e^2)^m} \tag{8.5.8}$$

and

$$I_2 = \int \frac{udu}{(u^2+e^2)^m}.$$

There is no problem at all in evaluating I_2. If we put $v = u^2 + e^2$, then $dv = 2udu$, and

$$I_2 = \frac{1}{2}\int \frac{dv}{v^m},$$

which equals

$$\frac{1}{2}\log|v| + C = \frac{1}{2}\log|u^2+e^2| + C \quad \text{if} \quad m = 1,$$

and

$$-\frac{1}{2(m-1)v^{m-1}} + C = -\frac{1}{2(m-1)(u^2+e^2)^{m-1}} + C \quad \text{if} \quad m > 1.$$

We may use substitution again to simplify I_1. Indeed, if we set $y = u/e$, we have

$$dy = \frac{1}{e}du, \quad \text{and} \quad u^2 + e^2 = e^2(y^2+1).$$

Consequently,

(8.5.9) $$I_1 = \frac{1}{e}\int \frac{dy}{(y^2+1)^m}.$$

As mentioned above, this will turn out to be given by $\frac{1}{e}\arctan y + C$ when $m = 1$.

It is left to discuss a **reduction process** which allows us to compute

$$J_m = \int \frac{dy}{(y^2+1)^m}$$

for $m > 1$.

Let us try the substitution

$$y = \tan t.$$

Since $\tan^2 t + 1 = \sec^2 t$, we have

$$(y^2+1)^m = \sec^{2m} t \quad \text{and} \quad dy = \sec^2 t dt.$$

Consequently, since $\frac{1}{\sec t} = \cos t$,

$$J_m = \int \cos^{2(m-1)} t\, dt.$$

We have already evaluated it in the previous chapter. Finally, to write the answer in the variable y, we will need to write t as $\arctan y$.

9 Inverse Functions, log, exp, arcsin, ...

9.1 Inverse functions

Suppose f is a function with domain X and image (or range) Y. By definition, given any x in X, there is a **unique** y in Y such that $f(x) = y$. But this definition of a function is not egalitarian, because it does not require that a *unique* number x be sent to y by f; so y is special but x is not. A really nice kind of a function is what one calls a **one-to-one** (or **injective**) function. By definition, f is such a function iff

(9.1.1) $$f(x) = f(x') \Longrightarrow x = x'.$$

In such a case, we can define an **inverse function** g with domain Y and range X, given by

(9.1.2) $$g(y) = x \quad \text{iff} \quad f(x) = y.$$

Clearly, when such an inverse function g exists, one has

(9.1.3) $$g \circ f = 1_X \quad \text{and} \quad f \circ g = i_Y,$$

where i_X, resp. 1_Y, denotes the **identity function** on X, resp. Y.

We will be concerned in this chapter with X, Y which are subsets of the real numbers.

Proposition 9.1.4 *Let f be a one-to-one function with domain $X \subset \mathbb{R}$ and range Y, with inverse g. Suppose in addition that f is differentiable at x with $f'(x) \neq 0$. Then g is differentiable at $y = f(x)$ and we have*

$$g'(y) = \frac{1}{f'(x)}$$

for all x in X with $y = f(x)$.

Note that if we know *a priori* that f and g are both differentiable, then this is easy to prove. Indeed, in that case their composite function $g \circ f$, which is the identity on X by (8.1.3), would be differentiable. By differentiating the identity

$$g(f(x)) = x$$

with respect to x, and applying the *chain rule*, we get

$$g'(f'(x)) \cdot f'(x) = 1,$$

because the derivative of x is 1. Done.

Proof. We have to compute the limit

$$L = \lim_{h \to 0} \frac{g(y+h) - g(y)}{h}.$$

Since f, g are one-to-one, we can find, for each small h, a small h' such that

$$y + h \;=\; f(x + h').$$

So

$$g(y + h) - g(y) \;=\; (x + h') - x \;=\; h'$$

and

$$h \;=\; (y + h) - y \;=\; f(x + h') - f(x).$$

Moreover, h' goes to 0 as h goes to 0 (and *vice-versa*). Hence

$$L \;=\; \lim_{h' \to 0} \frac{h'}{f(x + h') - f(x)},$$

which is the inverse of $f'(x)$. It makes sense because $f'(x)$ is by assumption non-zero.

□

Note that this proof shows that g is not differentiable at any point $y = f(x)$ if f' is zero at x.

It is helpful to note that many a function is not one-to-one in its maximal domain, but becomes one when restricted to a smaller domain. To give a simple example, the **squaring function**

$$f(x) \;=\; x^2$$

is defined everywhere on $\mathbb{R}$. But it is not one-to-one, because $f(a) = f(-a)$. However, if we restrict to the subset $\mathbb{R}_+$ of non-negative real numbers, f is one-to-one and so we may define its inverse to be the **square-root function**

$$g(y) \;=\; \sqrt{y}, \;\forall\, y \in \mathbb{R}_+.$$

Another example is provided by the sine function, which is periodic of period 2π and hence not one-to-one on $\mathbb{R}$. But it becomes one when restricted to $[-\pi/2, \pi/2]$.

9.2 The natural logarithm

For any $x > 0$, its natural logarithm is defined by the definite integral

$$\log x \;=\; \int_1^x \frac{dt}{t}. \tag{9.2.1}$$

Some write $\ln(x)$ instead, and some others write $\log_e x$. When $0 < x < 1$, this signifies the negative of integral of $\frac{1}{t}$ from x to 1. Consequently, $\log x$ is positive if $x > 0$, negative if $x < 0$ and equals 0 at $x = 1$.

Proposition 9.2.2

(a) $\log 1 = 0$.

(b) $\log x$ *is differentiable everywhere in its domain* $\mathbb{R}_+ = (0, \infty)$ *with derivative* $\frac{1}{x}$.

(c) *(***addition theorem***)* *For all* $x, y > 0$,

$$\log(xy) = \log x + \log y.$$

(d) $\log x$ *is a strictly increasing function.*

(e) $\log x$ *becomes unbounded in the positive direction when* x *goes to* ∞ *and it is unbounded in the negative direction when* x *goes to* 0.

(f) $\log x$ *is integrable on any finite subinterval of* $\mathbb{R}_+$, *and its indefinite integral is given by*

$$\int \log x \, dx = x \log x - x + C.$$

(g) $\log x$ *goes to* $-\infty$ *(resp.* ∞*) slowly when* x *goes to* 0 *(resp.* ∞*); more precisely,*

(i)
$$\lim_{x \to 0} x \log x = 0$$

and

(ii)
$$\lim_{x \to \infty} \frac{\log x}{x} = 0.$$

(h) *The* **improper integral** *of* $\log x$ *over* $(0, b]$ *exists for any* $b > 0$, *with*

$$\int_0^b \log x \, dx = b \log b - b.$$

Property (c) is very important, because it can be used to **transform multiplicative problems into additive ones**. This was the motivation for their introduction by **Napier** in 1616. Property (b) is also important. Indeed, if we assume only the properties (a),(c),(d),(e) for a function f on $\mathbb{R}_+$, there are lots of functions (*logarithms*) which satisfy these properties. But the situation becomes rigidified with a **unique solution** once one requires (b) as well. This is why log is called the **natural logarithm**. The other (unnatural) choices will be introduced towards the end of the next section.

Proof. (**a**): Since $\log 1 = 0$ and $\exp \circ \log$ is the identity function,

$$\exp(0) = \exp(\log(1)) = 1.$$

(**b**): For any x, the function $\frac{1}{t}$ is continuous on $[0, x]$, hence by the *First Fundamental Theorem of Calculus*, $\log x$ is differentiable at x with derivative $\frac{1}{x}$.
(**c**): Fix any $y > 0$ and consider the function of x defined by

$$\ell(x) = \log(xy) \quad \text{on} \quad \{x > 0\}.$$

Since it is the composite of the differentiable functions $x \to xy$ and $u \to \log u$, ℓ is also differentiable. Applying the *chain rule*, we get

$$\ell'(x) = y\left(\frac{1}{yx}\right) = \frac{1}{x}.$$

Thus ℓ and log both have the same derivative, and so their difference must be independent of x. Write

$$\ell(x) = \log x + c.$$

Evaluating at $x = 1$, and noting that $\log 1 = 0$ and $\lambda(1) = \log x$, we see that c must be $\log x$, proving the *addition formula*.
(**d**): As we saw in chapter 6, a differentiable function f is strictly increasing iff its derivative is positive everywhere. When $f(x) = \log x$, the derivative, as we saw above, is $1/x$, which is positive for $x > 0$. This proves (d).
(**e**): As $\log x$ is strictly increasing and since it vanishes at 1, its value at any number $x_0 > 1$, for instance at $x_0 = 2$, is positive. By the addition theorem and induction, we see that for any positive integer n,

$$\log(x_0^n) = n \log x_0. \tag{9.2.3}$$

Consequently, as n goes to ∞, $\log(x_0^n)$ goes to ∞ as well. This proves that $\log x$ is unbounded in the positive direction. For the negative direction, note that for any positive $x_1 < 1$, $\log x_1$ is negative (since $\log 1 = 0$ and $\log x$ is increasing). Applying (8,2,3) with x_0 replaced by x_1, we deduce that $\log(x_1^n)$ goes to $-\infty$ as n goes to ∞. Done.
(**f**) Since $\log x$ is continuous, it is integrable on any finite interval in $(0, \infty)$. Moreover, by *integration by parts*,

$$\int \log x \, dx = x \log x - \int x \frac{d}{dx}(\log x) dx.$$

The assertion (f) now follows since the expression on the right is $x \log x - x + C$.
(**g**): Put

$$L = \lim_{x \to 0} x \log x$$

and

$$u = \frac{1}{x}, \; f(u) = -\log\left(\frac{1}{u}\right), \quad \text{and} \quad g(u) = u.$$

Then we have

$$L = -\lim_{u \to \infty} \frac{f(u)}{g(u)}.$$

Then by (e), $f(u)$ and $g(u)$ approach ∞ as u goes to ∞, and both these functions are differentiable at any positive u. (All one needs is that they are differentiable for large u.) Since $u'(x) = -\frac{1}{x^2} = -u^2$ and $f(u) = -\log x$, we have by the *chain rule*,

$$f'(u) \;=\; \frac{df/dx}{du/dx} \;=\; \frac{-1/x}{-u^2} \;=\; \frac{1}{u}.$$

Since $g'(u) = 1$ for all u, we then get

$$\lim_{u\to\infty} \frac{f'(u)}{g'(u)} \;=\; 0.$$

So we may apply L'Hopital's rule (see the **Appendix** to this chapter), and conclude that

$$L \;=\; -\lim_{u\to\infty} \frac{f'(u)}{g'(u)} \;=\; 0,$$

giving (i). The proof of (ii) is similar and will be left for the reader to check.
(**h**): The improper integral of $\log x$ exists over $(0, b]$ iff the following limit exists:

$$L \;=\; \lim_{x\to 0} \int\limits_x^b \log t \, dt.$$

Thanks to (f), we have

$$\int\limits_x^b \log t \, dt \;=\; (t\log t - t)\big|_x^b \;=\; (b\log b - b) - x\log x + x.$$

To prove (h) we need to show that

$$\lim_{x\to 0} (x\log x - x) \;=\; 0,$$

which is a consequence of (g).

□

9.3 The exponential function

In view of the discussion in section 9.1, and the fact (parts (d), (e) of Proposition 9.2.2) that $\log x$ is a strictly increasing function with domain $\mathbb{R}_+$ and range $\mathbb{R}$, we can define the **exponential function**, $\exp(\mathbf{x})$ for short, to be the inverse function of $\log x$. Note that $\exp(x)$ has domain $\mathbb{R}$ and range $\mathbb{R}_+$.

Proposition 9.3.1

(a) $\exp(0) = 1$.

(b) $\exp(x)$ *is differentiable everywhere in* $\mathbb{R}$ *with derivative* $\exp(x)$.

(c) *(***addition theorem***)* *For all* $x, y \in \mathbb{R}$,

$$\exp(x+y) = \exp(x)\exp(y).$$

(d) $\exp(x)$ *is a strictly increasing function.*

(e) $\exp(x)$ *becomes unbounded in the positive direction when* x *goes to* ∞ *and it goes to* 0 *when* x *goes to* $-\infty$.

(f) $\exp(x)$ *is integrable on any finite subinterval of* $\mathbb{R}_+$*, and its indefinite integral is given by*

$$\int \exp(x)\, dx = \exp(x) + C.$$

(g) *As* x *goes to* ∞ *(resp.* $-\infty$*),* $\exp(x)$ *goes to* ∞ *(resp.* 0*) faster than any polynomial* $p(x)$ *goes to* ∞ *(resp.* $-\infty$*), i.e.,*

(i)
$$\lim_{x\to\infty} \frac{p(x)}{\exp(x)} = 0$$

and

(ii)
$$\lim_{x\to-\infty} p(x)\exp(x) = 0.$$

(h) *The* **improper integral** *of* $\exp(x)$ *over* $(-\infty, b]$ *exists for any* $b > 0$*, with*

$$\int_{-\infty}^{b} \exp(x)\, dx = \exp(b).$$

Proof. (**a**): This is immediate from the definition.
(**b**): Note that the derivative $\frac{1}{x}$ of $\log x$ is nowhere zero on its domain $\mathbb{R}_+$. So we may apply Proposition 9.1.4 with $f = \log$, $g = \exp$ and $y = \log x$ to get the everywhere differentiability of exp, with

$$\frac{d}{dy}\exp(y) = \frac{1}{(\log x)'} = \frac{1}{1/x} = x = \exp(y).$$

(**c**): Fix x, y in $\mathbb{R}$. Then we can find (unique) u, v in $\mathbb{R}_+$ such that

$$x = \log u \quad \text{and} \quad y = \log v.$$

Applying part (c) of Proposition 9.2.2, we then obtain

$$\exp(x+y) \;=\; \exp(\log u + \log v) \;=\; \exp(\log(uv)) \;=\; uv.$$

The assertion follows because $u = \exp(x)$ and $v = \exp(y)$.
(**d**): Let $x > y$ be arbitrary in $\mathbb{R}$. Write $x = \log u$, $y = \log v$ as above. Since log is strictly increasing, we need $u > v$. Since $u = \exp(x)$ and $v = \exp(v)$, we are done.
(**e**): Suppose $\exp(x)$ is bounded from above as x goes to infinity, i.e., suppose there is a positive number M such that $\exp(x) < M$ for all $x > 0$. Then, since log is a strictly increasing function, $x = \log(\exp(x))$ would be less than $\log(M)$ for all positive x, which is absurd. So $\exp(x)$ must be unbounded as x becomes large. To show that $\exp(x)$ approaches 0 as x goes to $-\infty$, we need to show that for any $\varepsilon > 0$, there exists $-T < 0$ such that

$$\exp(x) \;<\; \varepsilon \quad \text{whenever} \quad x < -T. \tag{9.3.2}$$

Applying the logarithm to both inequalities, writing $-T = \log u$ for a unique $u \in (0,1)$ and $\varepsilon' = \exp(\varepsilon)$, and using the fact that log is a one-to-one function, we see that (9.3.2) is equivalent to the statement

$$x \;<\; \varepsilon' \quad \text{whenever} \quad \log x < \log u, \tag{9.3.3}$$

which evidently holds by the properties of the logarithm. Hence $\exp(x)$ goes to 0 as x goes to $-\infty$.
(**f**) Since $\exp(x)$ is continuous, it is integrable on any finite subinterval of $\mathbb{R}$. In fact, since $\exp(x)$ is its own derivative, it is its own **primitive** as well, proving the assertion.
(**g**): To prove (i), it suffices to show that for any $j \geq 0$, the limit

$$L \;=\; \lim_{x\to\infty} x^j \exp(x)$$

exists and equals 0. When $j = 0$ this follows from (e), while the assertion for $j > 1$ follows from the case $j = 1$, which we will assume to be the case from here on. It then suffices to show that the function $\exp(x)/x$ goes to ∞ as x goes to ∞. By taking logarithms, this becomes equivalent to the statement that $x/\log x$ goes to ∞ as $\log x$, and hence x, goes to ∞, which is what we proved in our proof of the limit (i) of Proposition 9.2.2, part (g). (One can also apply, with care, L'Hopital's rule directly to the quotient $x/\exp(x)$ and thereby establish (i).) The proof of (ii) is similar and will be left for the reader to check.
(**h**): The improper integral of $\exp(x)$ exists over $(-\infty, b]$ iff the following limit exists:

$$L \;=\; \lim_{x\to-\infty} \int_x^b \exp(t)dt.$$

Thanks to (f), we have

$$\int_x^b \exp(t)\,dt \;=\; \exp(b) - \exp(x).$$

To prove (h) we need only show that

$$\lim_{x \to -\infty} \exp(x) \;=\; 0,$$

which is a consequence of (e).

$\square$

How many differentiable functions are there which are derivatives (and hence primitives) of themselves? This is answered by the following

Lemma 9.3.4 *Let f be any differentiable function on an open interval (a, b), possibly of infinite length, satisfying $f'(x) = f(x)$ for all x in (a, b). Then there exists a scalar c such that*

$$f(x) \;=\; c\exp(x) \;\forall\, x \in (a, b).$$

Moreover, if $f(0) = 1$, then $f(x)$ equals $\exp(x)$ *on this interval.*

Proof. Define a function h on (a, b) by

$$h(x) \;=\; \frac{f(x)}{\exp(x)},$$

which makes sense because exp never vanishes anywhere.

Since f and exp are differentiable, so is h, and by the *quotient rule* we have

$$h'(x) \;=\; \frac{f'(x)\exp(x) - f(x)\exp'(x)}{\exp(x)^2} \;=\; \frac{f(x)\exp(x) - f(x)\exp(x)}{\exp(x)^2} \;=\; 0,$$

because f and exp are derivatives of themselves. Therefore $h(x)$ must be a constant c, say. The Lemma now follows.

$\square$

Definition 9.3.5 *Put*

$$e \;=\; \exp(1).$$

Equivalently, e is the unique, positive real number whose natural logarithm is 1. It is not hard to see that, to a first approximation, $2 < e < 4$. One can do much better with some work, of course, and also show that e is **irrational**, even **transcendental**.

A natural question now arises. We have worked with the power function a^x before, for any positive real number a, which satisfies the same addition rule as $\exp x$, i.e.,

$$a^{x+y} \;=\; a^x \cdot a^y. \tag{9.3.6}$$

. What is the relationship between e^x and $\exp(x)$. The answer is very satisfying.

Proposition 9.3.7 *For all x in $\mathbb{R}$, we have*

$$\exp(x) \;=\; e^x.$$

We will for now accept the following result, which will be proved later in chapter 11:

Lemma 9.3.8 *We have*

$$\lim_{x\to 0} \frac{e^x - 1}{x} = 1.$$

Proof of Proposition 9.3.7. First we claim that e^x is its own derivative. Indeed, by (9.3.6),

$$\lim_{h\to 0} \frac{e^{x+h} - e^x}{h} = \lim_{h\to 0} e^x \left(\frac{e^h - 1}{h}\right).$$

The right hand side limit is e^x by Lemma 9.3.8, yielding the claim.

Next we apply Lemma 9.3.4 to deduce that e^x must be a multiple of $\exp(x)$. In fact, since $e^0 = 1$, e^x must equal $\exp(x)$.

□

Just like $\exp(x)$, the power function a^x is, for any positive number a, a strictly increasing function. One calls the inverse of a^x by the name **logarithm to the base** a and denotes it by $\log_a y$. Then $\log_a$ satisfies many of the properties of log, but its derivative is not $1/x$ (if $a \neq e$.

Proposition 9.3.9 *Fix $a > 0$. Then*

$$\frac{d}{dx}(a^x) = a^x \log a,$$

and

$$\frac{d}{dy}(\log_a y) = \frac{1}{y \log a}.$$

Proof. Note that

$$a^x = \exp(\log(a^x)) = e^{x \log a},$$

which is the composite of $x \to x \log a$ and $u \to e^u$. Applying the chain rule, we obtain

$$\frac{d}{dx}(a^x) = e^{x \log a} \cdot \log a = a^x \log a.$$

The formula for the derivative of $\log_a y$ follows from this and Proposition 9.1.4.

□

The basic **hyperbolic functions** are defined as follows:

$$\sinh x = \frac{e^x - e^{-x}}{2}, \ \cosh x = \frac{e^x + e^{-x}}{2}, \ \tanh(x) = \frac{\sinh x}{\cosh x}. \tag{9.3.10}$$

The other hyperbolic functions are given by the inverses of these.

9.4 arcsin, arccos, arctan, et al

The sine function, which is defined and differentiable everywhere on $\mathbb{R}$, is periodic with period 2π and is therefore not a one-to-one function. It however becomes one when the domain is restricted to $[-\pi/2, \pi/2]$. One calls the corresponding inverse function the **arcsine** function, denoted $\arcsin x$. Clearly, the domain of $\arcsin x$ is $[-1, 1]$ and the range is $[-\pi/2, \pi/2]$. Moreover, since the derivative $\cos x$ of $\sin x$ is non-zero, in fact positive, in $(-\pi/2, \pi/2)$, we may apply Proposition 9.1.4 and deduce that the arcsine function is differentiable on $(-\pi/2, \pi/2)$, with

$$\frac{d}{dy}(\arcsin y) \;=\; \frac{1}{\cos x} \;=\; \frac{1}{\sqrt{1-\sin^2 x}} \;=\; \frac{1}{1-y^2}, \tag{9.4.1}$$

where $y = \sin x$.

One notes similarly that $\cos x$, resp. $\tan x$, is one-to-one on $[0, \pi]$, resp. $(-\pi/2, \pi/2)$, and defines its inverse function $\arccos x$, resp. $\arctan x$, with domain $[-1, 1]$ and range $[0, \pi]$. Arguing as above, we see that $\arccos x$ is differentiable on $(0, \pi)$ and $\arctan x$ is differentiable on $(-\pi/2, \pi/2)$, with

$$\frac{d}{dy}(\arccos x) \;=\; -\frac{1}{\sqrt{1-y^2}} \tag{9.4.2}$$

and

$$\frac{d}{dy}(\arctan x) \;=\; \frac{1}{1+y^2} \tag{9.4.3}$$

Consequently, we obtain

$$\int \frac{dy}{1+y^2} \;=\; \arctan y + C \tag{9.4.4}$$

and

$$\int \frac{dy}{\sqrt{1-y^2}} \;=\; \arcsin y + C. \tag{9.4.5}$$

This incidentally brings to a close our quest to integrate arbitrary rational functions, which we began in the previous chapter, where we reduced the problem to the evaluation of the integral on the left of (9.4.4).

It should be notes that it is a miracle that we can evaluate the reciprocal of the square root of $1 - y^2$ (for $-1 \leq y \leq 1$) in terms of $\arcsin y$. If one tried to integrate $\frac{1}{\sqrt{f(y)}}$ for a polynomial f of degree $n > 2$, the problem becomes forbiddingly difficult. Even for $n = 3$, one needs to use **elliptic functions**.

9.5 A useful substitution

A very useful substitution to deal with trigonometric integrals is to set

$$u = \tan(x/2), \tag{9.5.1}$$

which implies that

$$x = 2\arctan u.$$

Note that

$$\frac{du}{dx} = \frac{1}{2}\sec^2(x/2)\,\frac{1}{2}(1+\tan^2(x/2)) = \frac{1+u^2}{2}, \tag{9.5.2}$$

$$dx = \frac{2}{1+u^2}du, \tag{9.5.3}$$

$$\sin x = 2\sin(x/2)\cos(x/2) = 2\frac{\tan(x/2)}{\sec^2(x/2)} = \frac{2u}{1+u^2}, \tag{9.5.4}$$

and since $\cos^2(x/2) + \sin^2(x/2) = 1$,

$$\cos x = \frac{\cos^2(x/2) - \sin^2(x/2)}{\cos^2(x/2) + \sin^2(x/2)} = \frac{1-\tan^2(x/2)}{1+\tan^2(x/2)} = \frac{1-u^2}{1+u^2}. \tag{9.5.5}$$

For example, suppose we have to integrate

$$I = \int \frac{dx}{1-\sin x}.$$

Using the substitution above, which is justifiable here, we get

$$I = \int \frac{1}{1-2u/(1+u^2)}\frac{2}{1+u^2}du = \int 2\frac{du}{1-2u+u^2}.$$

Since

$$\frac{1}{1-2u+u^2} = \frac{1}{(1-u)^2} = \frac{d}{du}\left(\frac{1}{1-u}\right),$$

we get

$$I = \int d\left(\frac{1}{1-u}\right) = \frac{1}{1-u} + C = \frac{1}{1-\tan(x/2)} + C.$$

9.6 Appendix: L'Hopital's Rule

It appears that the most popular mathematician for Calculus students is **Marquis de L'Hopital**, who was prolific during the end of the seventeenth century. Everyone likes to use his **rule**, but two things must be taken due note of. The first is that, as with any other theorem, one has to make sure that all the hypotheses hold before applying it. The second is a bit more subtle. One should not use it when it leads to a circular reasoning, for example when the numerator or the denominator of the limit L in question, which goes to 0 or ∞ as the case might be, is differentiable as needed, but to prove it one needs the limit L to exist in the first place. Here is an example to illustrate this point. Consider the following two statements:

(I) The function e^x is differentiable with derivative e^x.

(II) The limit

$$L = \lim_{t\to 0} \frac{e^t - 1}{t}$$

equals 1.

When presented with the limit L one is tempted to prove that it is 1 by using L'Hopital's rule. Indeed, if we accept that it is applicable here, then, since e^t is by (I) differentiable with derivative e^t, and since the limit

$$L_1 = \lim_{t\to 0} \frac{e^t}{1}$$

equals $e^0 = 1$, one thinks that the problem is solved. But not so fast! This method assumes (I) and how does one prove it? Well, one has to show the following:

$$\lim_{h\to 0} \frac{e^{x+h} - e^x}{h} = e^x.$$

But since e^{x+h} is $e^x e^h$, one has to show that

$$\lim_{h\to 0} \frac{e^h - 1}{h} = 1,$$

which is the assertion (II). So (I) and (II) are equivalent and one cannot prove one using the other, unless one has found a different way to prove one of them. So if one uses the L'Hopital's rule to evaluate L, one has to show why e^x is differentiable without using L as a tool, which can be done. Of course one defines the exponential function $\exp(x)$ as the inverse function of the logarithm and the fact that the derivative of $\log x$ is $1/x$ implies, as we saw earlier, that $\exp(x)$ is differentiable with derivative $\exp(x)$. But we then have to show, by another method, that e^x is the same as $\exp(x)$. If one is not careful one will be drawn into a delicate spider web.

This is not to scare you into not using L'Hopital's rule. Just make sure before using it that you can satisfy the hypotheses and that there are no circular arguments. Make sure,

in particular, that the numerator and the denominator of the limit L can be shown to be differentiable without using L. Indeed, the most important thing one has to learn in Ma 1a is to **think logically**.

Without further ado, let us now present the rule of L'Hopital.

Theorem (L'Hopital's rule) *Consider a limit of the form*

$$L = \lim_{x \to a} \frac{f(x)}{g(x)},$$

where a is either 0 or ∞ or $-\infty$ or just any (finite) non-zero real number, and f, g are differentiable functions at all real numbers x with $|x - a|$ sufficiently small. Suppose both and f and g approach 0, or both approach ∞, or both tend to $-\infty$, as x goes to a. Then L exists **if** *the* **limit quotient of the derivatives***, namely*

$$L' = \lim_{x \to a} \frac{f'(x)}{g'(x)}$$

exists. Moreover, when L' exists, L equals L'.

To prove this we need the following generalization of the usual *Mean value theorem*:

Cauchy's Mean Value Theorem *Suppose f, g are functions which are continuous on $[a, b]$ and differentiable on (a, b). Then there is a number c in (a, b) such that*

$$(f(b) - f(a))g'(c) = (g(b) - g(a))f'(c).$$

Note that when $g(x) = x$, we recover the statement of the usual Mean Value Theorem.

Proof of the Cauchy Mean Value Theorem. Put

$$u(x) = f(x)(g(b) - g(a)) - g(x)(f(b) - f(a)),$$

which is continuous on $[a, b]$ and differentiable on (a, b) because f and g are. Moreover,

$$u(a) = f(a)(g(b) - g(a)) - g(a)(f(b) - f(a)) = f(a)g(b) - g(a)f(b)$$

and

$$u(b) = f(b)(g(b) - g(a)) - g(b)(f(b) - f(a)) = -g(a)f(b) + f(a)g(b).$$

Hence

$$u(a) = u(b).$$

So we may apply Rolle's theorem and conclude that there exists a c in (a, b) such that $u'(c) = 0$. In other words,

$$f'(c)(g(b) - g(a)) - g'(c)(f(b) - f(a)) = 0,$$

which is what we wanted to prove.

□

Proof of L'Hopital's Rule. We will prove this in the case when $a = \infty$, with $f(x), g(x)$ both approaching ∞ as x approaches ∞. The other cases require only very slight modifications and will be left as exercises for the interested reader.

By hypothesis, $f(x)$ and $g(x)$ are defined and differentiable for large enough x. Suppose the limit L' exists. We have to show that L also exists, and prove that in fact $L = L'$.

The existence of L' as a (finite) real number implies that for every $\varepsilon > 0$, there is some $b > 0$ such that for all $x > b$,

$$(A1) \qquad \left| \frac{f'(x)}{g'(x)} - L' \right| < \varepsilon.$$

Since $g(x)$ goes to ∞ as $x \to \infty$, we may choose b large enough so that $g(x) \neq g(b)$ for all $x > b$.

By applying the Cauchy Mean Value Theorem to $[b, x]$, we get

$$(A2) \qquad \frac{f'(c)}{g'(c)} = \frac{f(x) - f(b)}{g(x) - g(b)},$$

for some c in (b, x). Combining with (A1), we then get

$$(A3) \qquad \left| \frac{f(x) - f(b)}{g(x) - g(b)} - L' \right| < \varepsilon.$$

In other words,

$$(A4). \qquad \lim_{x\to\infty} \frac{f(x) - f(b)}{g(x) - g(b)} = L.$$

Now we are almost there. To finish, note that since $f(x) \to \infty$ as $x \to \infty$,

$$(A5) \qquad \lim_{x\to\infty} \frac{f(x)}{f(x) - f(b)} = 1.$$

Similarly,

$$(A6) \qquad \lim_{x\to\infty} \frac{g(x) - g(b)}{g(x)} = 1.$$

The assertion now follows by combining (A4), (A5) and (A6).

□

10 Taylor's theorem, Polynomial approximations

Polynomials are the nicest possible functions. They are easy to differentiate and integrate, which is also true of the basic trigonometric functions, but more importantly, polynomials can be evaluated at any point, which is not true for general functions. So what one does in practice is to approximate any function f of interest by polynomials. When the approximation is done by *linear polynomials*, then it is called a *linear approximation*, which pictorially corresponds to *linearizing* the graph of f. It turns out that the more times one can differentiate f, the higher is the degree of the polynomial one can approximate it with, and more importantly, the better the approximation becomes, as one sees it intuitively. There is only one main theorem here, due to Taylor, but it is omnipresent in all the mathematical sciences, with a number of ramifications, and should be understood precisely.

10.1 Taylor polynomials

Suppose f is an N-times differentiable function on an open interval I. Fix any point a in I. Then for any non-negative integer $n \leq N$, the **nth Taylor polynomial of f at $x = a$** is given by

(10.1.1) $$p_n(f(x); a) = \sum_{j=0}^{n} \frac{f^{(j)}(a)}{j!}(x-a)^j,$$

where $f^{(j)}(a)$ denotes the jth derivative of f at a. By convention, $f^{(0)}(a)$ just denotes $f(a)$. (f is the 0th derivative of itself!)

The coefficients $\frac{f^{(j)}(a)}{j!}$ are called the **Taylor coefficients of f at a.**

The definition has been rigged so that the following holds:

Lemma 10.1.2 *Suppose f is itself a polynomial, i.e.,*

$$f(x) = a_0 + a_1 x + \ldots + a_m x^m,$$

for some integer $m \geq 0$. Then f is infinitely differentiable (which means it can be differentiated any number of times), and

$$p_n(f(x); 0) = \begin{cases} a_0 + a_1 x + \ldots + a_n x^n, & \text{if} \quad n < m \\ a_0 + a_1 x + \ldots + a_m x^m, & \text{if} \quad n \geq m \end{cases}$$

Proof. Clearly, f is differentiable any number of times and moreover, $f^{(n)}(x)$ vanishes if $n > m$. So we have only to show that for $n \leq m$,

(10.1.3) $$f^{(n)}(0) = n! a_n.$$

When $m = 0$ this is clear. So let $m > 0$ and assume by induction that (10.1.3) holds for all polynomials of degree $m-1$ and $n \le m-1$. Define a polynomial $g(x)$ by the formula

$$f(x) \;=\; a_0 + xg(x).$$

Then

$$g(x) \;=\; \sum_{j=0}^{m-1} a_{j+1}x^j$$

and by the inductive hypothesis,

$$g^{(n)}(0) \;=\; n!a_{n+1} \tag{10.1.4}$$

for all non-negative $n \le m-1$. But by the *product rule*,

$$f'(x) \;=\; g(x) + xg'(x),\; f''(x) \;=\; 2g'(x) + xg''(x), \ldots$$

By induction, we get

$$f^{(n)}(x) \;=\; ng^{(n-1)}(x) + xg^{(n)}(x),$$

so that

$$f^{(n)}(0) \;=\; ng^{(n-1)}(0)\;\forall\, n \le m, n \ge 1. \tag{10.1.5}$$

The identity (10.1.3), and hence the Lemma, now follow by combining (10.1.4) and (10.1.5). □

Lemma 10.1.6 (Linearity) *Let f, g be n-times differentiable at a, and let α, β be arbitrary scalars. Then*

$$p_n(\alpha f(x) + \beta g(x); a) \;=\; \alpha p_n(f(x); a) + \beta p_n(g(x); a).$$

This is easy to prove because the derivative is linear. In particular, we have

$$\frac{(\alpha f + \beta g)^{(j)}(a)}{j!} \;=\; \alpha\frac{f^{(j)}(a)}{j!} + \beta\frac{g^{(j)}(a)}{j!}.$$

It is helpful to look at some **examples**:

(**1**): Let

$$f(x) \;=\; \sin x,$$

which is infinitely differentiable, with

$$f'(x) \;=\; \cos x,\; f''(x) \;=\; -\sin x \;=\; -f(x).$$

Thus

$$f^{(n)}(x) \;=\; \begin{cases} (-1)^k \sin x, & \text{if} \quad n = 2k \\ (-1)^k \cos x, & \text{if} \quad n = 2k+1 \end{cases}$$

Since $\sin 0 = 0$ and $\cos 0 = 1$, the Taylor polynomials of $\sin x$ are given by

$$p_0(\sin x;0) \;=\; 0,\;\; p_1(\sin x;0) \;= p_2(\sin x;0) \;=\; x,\;\; p_3(\sin x;0) \;=\; p_4(\sin x;0) \;=\; x - \frac{x^3}{6}, \ldots$$

More generally, for any positive integer k,

$$p_{2k-1}(\sin x;0) \;=\; p_{2k}(\sin x;0) \;=\; x - \frac{x^3}{3!} + \frac{x^5}{5!} - \ldots - (-1)^k \frac{x^{2k-1}}{(2k-1)!}. \tag{10.1.7}$$

(**2**): Put

$$f(x) \;=\; \log x.$$

This function is not defined at 0, so we need to choose another point to evaluate the derivatives, and the easiest one is

$$a \;=\; 1.$$

We have

$$f'(x) \;=\; \frac{1}{x},\; f''(x) \;=\; -\frac{1}{x^2},\; f'''(x) \;=\; \frac{2!}{x^3}, \ldots$$

By induction, we have for any $n \geq 1$,

$$f^{(n)}(x) \;=\; (-1)^{n+1}\frac{(n-1)!}{x^n}.$$

So the nth Taylor coefficient is

$$\frac{f^{(n)}(1)}{n!} \;=\; (-1)^{n+1}\frac{1}{n},$$

where we have used the simple fact that $n!$ is n times $(n-1)!$. Consequently, since $\log 1 = 0$, the nth Taylor polynomial of $\log x$ is given by

$$p_n(\log x;1) \;=\; x - \frac{x^2}{2} + \ldots + (-1)^{n+1}\frac{x^n}{n}. \tag{10.1.8}$$

(**3**): Consider

$$g(x) \;=\; \frac{1}{x}.$$

One has, for every $n \geq 0$,

$$g^{(n)}(x) \;=\; f^{(n+1)}(x),$$

where $f(x)$ is $\log x$. Thus for any $a > 0$,

$$\frac{g^{(n)}(1)}{n!} \;=\; (n+1)\frac{f^{(n+1)}(a)}{(n+1)!}. \tag{10.1.9}$$

As a consequence the Taylor polynomials of g at $a = 1$ are determinable from those of f. Let us make this idea precise.

Lemma 10.1.10 *Let f be a which is n times differentiable around a, with*

$$p_n(f(x);a)) \;=\; a_0 + a_1(x-a) + \ldots + a_n(x-a)^n.$$

Then

$$p_{n-1}(f'(x);a) \;=\; a_1 + 2a_2x + \ldots + na_n(x-a)^{n-1}.$$

Moreover, if ϕ is a primitive of f around a,

$$p_n(\phi(x);a) \;=\; \phi(a) + a_0(x-a) + a_1\frac{(x-a)^2}{2} + \ldots + a_{n-1}\frac{(x-a)^n}{n}.$$

The proof is immediate from the definition of Taylor polynomials.

For a general f, even for such a simple function like $\frac{1}{1+x^2}$, it is painful to work out the Taylor polynomials from scratch. One needs a better way to find them, and this will be accomplished in the next section.

10.2 Approximation to order n

Definition 10.2.1 *Let f, g be n times differentiable functions at a. We will say that f and g* **agree up to order** *n at a iff we have*

$$\lim_{x\to a}\frac{f(x)-g(x)}{(x-a)^n} \;=\; 0.$$

If g is a polynomial agreeing with f (or equalling f, as some would say) up to order n, then we would call g a **polynomial approximation of** $f(x)$ **to oder** n **at** $x = a$. The immediate question which arises is whether the nth Taylor polynomial of f is a polynomial approximation to order n. The answer turns out to be **Yes**, but even more importantly, the Taylor polynomial is the only one which has this property. Here is the complete statement!

Proposition 10.2.2 *Let f be n times differentiable at a. Then*

(i) $p_n(f(x);a)$ is a polynomial approximation of f to order n;

(ii) If $q(x)$ is any polynomial in $(x-a)$ of degree $\leq n$ which agrees with f up to order n, then $q(x) = p_n(f(x);a)$.

Proof. (**i**): Put

(10.2.3)
$$g(x) \;=\; p_{n-1}(f(x);a) \quad \text{and} \quad h(x) \;=\; (x-a)^n.$$

Then by definition,

$$p_n(f(x);a) \;=\; g(x) + \frac{f^{(n)}(a)}{n!}h(x).$$

Hence

$$\frac{f(x) - p_n(f(x);a)}{(x-a)^n} \;=\; \frac{f(x)-g(x)}{h(x)} - \frac{f^{(n)}(a)}{n!}.$$

So it suffices to prove the following:

$$\lim_{x\to a} \frac{f(x)-g(x)}{h(x)} \;=\; \frac{f^{(n)}(a)}{n!}. \tag{10.2.4}$$

Applying Lemma 10.1.2, we get

$$g^{(j)}(a) \;=\; f^{(j)}(a)\ \forall\, j < n. \tag{10.2.5}$$

Since g is a polynomial of degree $\leq n-1$, its $(n-1)$th derivative is a constant; so

$$g^{(n-1)}(x) \;=\; g^{(n-1)}(a). \tag{10.2.6}$$

Also,

$$h^{(j)}(x) \;=\; \frac{n!(x-a)^{n-j}}{(n-j)!}. \tag{10.2.7}$$

It follows from (10.2.5) and (10.2.7) that for every $j < n-1$,

$$\lim_{x\to a} f^{(j)}(x) - g^{(j)}(x) \;=\; \frac{f^{(j)}(a)-g^{(j)}(a)}{h^{(j)}(a)} \;=\; 0 \tag{10.2.8}$$

and

$$\lim_{x\to a} h^{(j)}(x) \;=\; h^{(j)}(a) \;=\; 0.$$

On the other hand, by (10.2.6) and (10.2.7),

$$\lim_{x\to a} \frac{f^{(n-1)}(x)-g^{(n-1)}(x)}{h^{(n-1)}(x)} \;=\; \lim_{x\to a} \frac{f^{(n-1)}(x)-f^{(n-1)}(a)}{n!(x-a)} \;=\; \frac{f^{(n)}(a)}{n!}. \tag{10.2.9}$$

In view of (10.2.8) and (10.2.9), we may apply L'Hopital's rule (see the Appendix to chapter 9) and deduce (10.2.4), which also proves part (i) of the Proposition.

(**ii**): By hypothesis, $q(x)$ approximates $f(x)$ to order n at a. By part (i), the Taylor polynomial $p_n(f(x);a)$ does the same thing. It follows, since the limit of a sum is the sum of the limits, that $q(x)$ and $p_n(x)$ agree up to order n. Put

$$u(x) \;=\; p_n(f(x);a) - q(x),$$

which is a polynomial of degree $\leq n$ and satisfies

$$\lim_{x\to a} \frac{u(x)}{(x-a)^n} \;=\; 0.$$

This implies in particular that

$$\lim_{x\to a} \frac{u(x)}{(x-a)^j} \;=\; 0 \;\forall\, j \le n. \tag{10.2.10}$$

On the other hand, applying the *Euclidean algorithm* repeatedly, relative to the divisor $(x-a)$, we can find, as we did in the chapter on*partial fractions*, numbers $c_0, \ldots, c_n$ such that

$$u(x) \;=\; c_0 + c_1(x-a) + \ldots + c_n(x-a)^n.$$

It is then immediate that for any $j \le n$,

$$\lim_{x\to a} \frac{u(x)}{(x-a)^j} \;=\; c_j.$$

In view of (10.2.10), this means that every coefficient c_j is zero. Thus the polynomial $u(x)$ is identically zero.

□

Now let us apply this to compute the Taylor polynomials of

$$\phi(x) \;=\; \arctan x \tag{10.2.11}$$

at $a = 0$, where ϕ takes the value 0. (You may try as an educational exercise to compute directly with $\phi(x)$, and you will learn why this Proposition is helpful.)

Recall that ϕ is a primitive of

$$f(x) \;=\; \frac{1}{1+x^2} \tag{10.2.12}$$

for all x in the domain of $\arctan x$, namely the open interval $(-\pi/2, \pi/2)$. Also, f is infinitely differentiable everywhere.

We will compute the Taylor polynomials of $f(x)$ at 0 by the following trick. For each $n \ge 1$, look at the polynomial

$$g_n(x) \;=\; 1 - x^2 + \ldots + (-1)^n x^{2n}.$$

It is a *geometric sum* and so we can reexpress it as

$$g_n(x) \;=\; \frac{1 + x^{2(n+1)}}{1+x^2} \;=\; f(x) + \frac{x^{2n+2}}{1+x^2}.$$

Consequently, for $r = 2n, 2n+1$

$$\lim_{x\to 0} \frac{f(x) - g_n(x)}{x^r} \;=\; \lim_{x\to a} \frac{x^{2n+2-r}}{1+x^2} \;=\; 0.$$

Thus $g_n(x)$ approximates $f(x)$ to order $2n$ and $2n+1$, so it must, by the Proposition above, equal the Taylor polynomials $p_{2n}(f(x);0)$ and $p_{2n+1}(f(x);0)$. Thus for any $n \ge 0$,

$$p_{2n}(\frac{1}{x};0) \;=\; p_{2n+1}(\frac{1}{x};0) \;=\; 1 - x^2 + \ldots + (-1)^n x^{2n}. \tag{10.2.13}$$

Applying Lemma 10.1.10, we then deduce that for all $n \geq 0$,
(10.2.14)
$$p_{2n+1}(\arctan x; 0) \;=\; p_{2n+2}(\arctan x; 0) \;=\; p_0(\arctan x; 0) + x - \frac{x^3}{3} + \ldots + (-1)^n \frac{x^{2n+1}}{2n+1}.$$

10.3 Taylor's Remainder Formula

Once one has looked at the Taylor polynomials $p_n(f; a)$ of a sufficiently differentiable function f at a point a, the natural question which arises immediately is how close an approximation to f does one get this way. To be precise, define the **nth remainder of f at a** to be

$$R_n(f(x), a) \;=\; f(x) - p_n(f(x); a). \tag{10.3.1}$$

A very precise answer to this question wa supplied by Taylor. Here it is!

Theorem 10.3.2 *Let $n \geq 0$, $a < x \in \mathbb{R}$, and f an $(n+1)$-times differentiable function on an open interval containing $[a, x]$. Then we have the following:*

(a)
$$R_n(f(x); a) \;=\; \frac{f^{(n+1}(c)}{(n+1)!}(x-a)^{n+1}$$
for some c in (a, x).

(b) If $f^{(n+1)}$ is integrable on $[a, x]$, then
$$R_n(f(x); a) \;=\; \frac{1}{n!}\int_a^x f^{(n+1)}(u)(x-u)^n du.$$

Corollary 10.3.3 *Let f be $(n+1)$-times differentiable on $[a, x]$. Suppose there are numbers m, M such that*
$$m \;\leq\; f^{(n+1)}(u) \;\leq\; M$$
for **all** *u in $[a, x]$. Then we have*

$$m\frac{(x-a)^{n+1}}{(n+1)!} \;\leq\; R_n(f)x); a) \;\leq\; M\frac{(x-a)^{n+1}}{(n+1)!}. \tag{i}$$

In particular, if $C = \max\{|m|, |M|\}$,

$$|R_n(f)x); a)| \;\leq\; C\frac{(x-a)^{n+1}}{(n+1)!}. \tag{ii}$$

Completely analogous assertions hold when $x < a$, in which case one should replace $[a, x]$, everywhere in the Theorem and Corollary with $[x, a]$, resp. $a - x$.

Let us first look at the **example** of the exponential function. We know that

$$f(u) \;=\; \exp(u)$$

is infinitely differentiable on all of $\mathbb{R}$ with $f'(u) = f(u)$. Moreover, since e^u is an increasing function with $e^0 = 1$, we get, for $x > 0$, $u \in [1, x]$ and $n \geq 0$,

$$1 \;\leq\; f^{(n+1)}(u) \;\leq\; e^x.$$

Consequently, by corollary 10.3.3, we have

$$\frac{x^{n+1}}{(n+1)!} \;\leq\; R_n(e^x; 0) \;\leq\; e^x \frac{x^{(n+1)}}{(n+1)!}. \tag{10.3.4}$$

Suppose we want to evaluate e to within an error of 10^{-4}. Then what we have to do is the following. Putting $x = 1$ in (8.3.4), and remembering the crude estimate that e is less than 3, we obtain

$$\frac{1}{(n+1)!} \;\leq\; R_n \;=\; R_n(e; 0) \;\leq\; \frac{3}{(n+1)!}. \tag{10.3.5}$$

Find the smallest n for which

$$\frac{3}{(n+1)!} \;<\; 10^{-4}.$$

Direct computation shows that

$$\frac{3}{7!} \;=\; \frac{3}{5040} \;=\; \frac{1}{1680} \;>\; 10^{-4}$$

and

$$\frac{3}{8!} \;=\; \frac{3}{40320} \;=\; \frac{1}{13440} \;<\; 10^{-4}.$$

So we take $n = 7$, and the error will be less than 10^{-4} if we approximate e by

$$p_7(e; 0) \;=\; 1 + 1 + \frac{1}{2!} + \frac{1}{3!} + \frac{1}{4!} + \frac{1}{5!} + \frac{1}{6!} + \frac{1}{7!},$$

which is

$$\frac{13700}{5040} \;=\; 2.7182539\ldots$$

The first four places after the decimal point are correct, as they should be. But at the fifth place the digit should be 8 instead of 5, and to get that one has to go to the n (namely 8) which makes R_n less than 10^{-5}.

The remainder formula applied to the functions $\sin x$ and $\cos x$ yields very similar estimates for the remainder. To be precise, we use the fact that the Taylor polynomials of $\sin x$, resp. $\cos x$, at $x = 0$, have only odd, resp. even, degree terms, and obtain the following:

$$\sin x \;=\; x - \frac{x^3}{3!} + \frac{x^5}{5!} - \ldots + (-1)^m \frac{x^{2m+1}}{(2m+1)!} + R_{2m+1}(\sin x; 0), \tag{10.3.5}$$

with

$$|R_{2m+1}(\sin x; 0)| \leq \frac{|x|^{2m+3}}{(2m+3)!};$$

and

$$(10.3.6) \qquad \cos x \;=\; 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \ldots + (-1)^m \frac{x^{2m}}{(2m)!} + R_{2m}(\cos x; 0),$$

with

$$|R_{2m+1}(\cos x; 0)| \leq \frac{|x|^{2m+2}}{(2m+2)!}.$$

It is a simple exercise to approximate numbers like $\sin 1$ or $\cos(1/2)$ to any number of decimel places.

Taylor's formula is not very useful, however, for estimating the remainders of functions f for which it is hard to get a nice expression for $f^{(n+1)}(u)$. A very **important example** illustrating this phenomenon is the function

$$f(x) \;=\; \arctan x, \;\; x \in (-\pi/2, \pi/2).$$

So what does one do? After some reflection, one remembers the method by which one found the Taylor polynomials of this functions. Luckily, this method also leads to a good estimate for the remainder. Let us see how.

Recall that

$$\frac{d}{dx}(\arctan x) \;=\; \frac{1}{1+x^2},$$

and that

$$\frac{1 - (-1)^m x^{2m+2}}{1+x^2} \;=\; 1 - x^2 + x^4 - \ldots + (-1)^m x^{2m}.$$

The second formula can be rewritten as

$$\frac{1}{1+x^2} \;=\; 1 - x^2 + x^4 - \ldots + (-1)^m \frac{x^{2m+2}}{1+x^2}.$$

Integrating this expression and using the fact that $\arctan 0 = 0$, we get by the fundamental theorem of Calculus,

$$(10.3.8 - i) \qquad \arctan x \;=\; x - \frac{x^3}{3} + \frac{x^5}{5} - \ldots + (-1)^m \frac{x^{2m+1}}{2m+1} + R_{2m+1}(\arctan x; 0),$$

where

$$R_{2m+1}(\arctan x; 0) \;=\; (-1)^m \int_0^x \frac{u^{2m+3}}{1+u^2}.$$

I am using here the fact that we have already seen that the polynomial $x - \frac{x^3}{3} + \frac{x^5}{5} - \ldots + (-1)^m \frac{x^{2m+1}}{2m+1}$ is the Taylor polynomial by the criterion given by part (ii) of Proposition 10.2.2. Suppose $x > 0$. Then

$$\frac{u^{2m+3}}{1+u^2} \;\leq\; u^{2m+3} \;\forall\, u \in [0, x],$$

in fact with equality only for $u = 0$, and hence

$$|R_{2m+1}(\arctan x; 0)| \le \int_0^x u^{2m+3} du.$$

Since the integral of u^{2m+3} is $u^{2m+4}/(2m+4)$, we get the desired bound

$$(10.3.8-ii) \qquad |R_{2m+1}(\arctan x; 0)| < \frac{x^{2m+3}}{2m+3}.$$

Note that for any fixed $x > 0$, this expression goes to 0 as m goes to ∞. By taking $x = 1$, and letting $m \to \infty$, one gets the **Leibniz formula**:

$$\frac{\pi}{4} = \arctan 1 = 1 - \frac{1}{3} + \frac{1}{5} - \ldots + (-1)^m \frac{1}{2m+1} + \ldots$$

This is no doubt a beautiful formula, but it is not quite useful for computations, because $1/m$ goes to 0 rather slowly, at least compared to $1/n!$, which is what one had for the exponential or the sine function. The silver lining is that while (10.3.8-ii) is not decreasing fast for $x = 1$, it converges faster when x is small. To exploit this, one appeals to the **addition theorem for the arctangent**, namely

$$(10.3.9) \qquad \arctan x + \arctan y = \arctan\left(\frac{x+y}{1-xy}\right),$$

which follows by applying the inverse function arctan to the addition theorem for the tangent function (with $x = \tan u, y = \tan v$):

$$\tan(u+v) = \frac{\tan u + \tan v}{1 - \tan u \tan v}.$$

¿From this one can derive, for example the following identities:

$$\frac{\pi}{4} = \arctan 1 = \arctan(1/2) + \arctan(1/3)$$

and

$$\frac{\pi}{4} = \arctan 1 = 4\arctan(1/5) - \arctan(1/239).$$

The second formula, proved by Machin in 1706, can be used to find the first five or six decimal places of π very fast. (Of course Mathematica or Maple can spew out the first 10,000 digits in a few seconds, but the methods used there are very sophisticated and appeal to formulas involving elliptic functions.)

Proof of Theorem 10.3.2, For every u in $[a, x]$, we have

$$(10.3.10) \qquad f(x) = p_n(f(x); u) + R_n(f(x); u),$$

where

$$p_n(f(x);u) \;=\; f(u) + f'(u)(x-u) + \frac{f''(u)}{2!}(x-u)^2 + \ldots + \frac{f^{(n)}(u)}{n!}(x-u)^n. \tag{10.3.11}$$

Note that
(10.3.12)

$$\frac{d}{du}p_n(f(x);u) \;=\; f'(u)+(-f'(u)+f''(u)(x-u))+\left(-f''(u)(x-u)+\frac{f^{(3)}(u)}{2!}(x-u)^2\right)+\ldots+\left(-\frac{-f^{(n)}(u)}{(n-1)!}\right.$$

Differentiating both sides of (10.3.10) and making use of (10.3.12), we obtain, for every u in $[a,x]$,

$$0 \;=\; \frac{f^{(n+1)}(u)}{n!}(x-u)^n + \frac{d}{du}R_n(f(x);u). \tag{10.3.13}$$

The function $R_n(f(x);u)$ is continuous on $[a,x]$ and differentiable, because $f(x)$ and $p_n(f(x);u)$ are, on (a,x). Of course the polynomial function $\phi(u) = (x-u)^{n+1}$ has the same properties. So we may apply the Cauchy Mean Value Theorem (see the Appendix to chapter 9) to $R_n(f(x);u)$ and $\phi(u)$ and get a number c in (a,x) such that

$$\frac{d\phi}{du}(c)(R_n(f(x);x) - R_n(f(x);a)) \;=\; \frac{d}{du}R_n(f(x);u)(c)(\phi(x)-\phi(a)). \tag{10.3.15}$$

By (10.3.13),

$$\frac{d}{du}R_n(f(x);u)(c) \;=\; -\frac{f^{(n+1)}(c)}{n!}(x-c)^n. \tag{10.3.16}$$

And

$$\frac{d\phi}{du}(c) \;=\; -(n+1)(x-c)^n. \tag{10.3.17}$$

Combining (10.3.15), (10.3.16) and (10.3.17), cancelling $-(x-u)^n$, and dividing by $(n+1)$, we obtain the formula (i). (This particular form of the remainder was in fact derived by Lagrange.)

Now suppose $f^{(n+1)}$ is integrable on $[a,x]$. Then applying the fundamental theorem of Calculus, and remembering that $R_n(f(x);x) = 0$, we get

$$-R_n(f(x);a) \;=\; \int_a^x \left(\frac{d}{du}R_n(f(x);u)\right) du,$$

whose right hand side expression is, by (10.3.13),

$$-\int_a^x \frac{f^{(n+1)}(u)}{n!}(x-u)^n du.$$

Hence we get (ii).

□

10.4 The irrationality of e

Now let us prove that e is irrational. It is even transcendental, but that is much harder to prove.

Suppose $e = p/q$ for some positive integers p, q. Choose an integer $n > 3$ which is greater than q. Using (10.3.4) and (10.3.5), we get

$$e = \frac{p}{q} = 1 + 1 + \frac{1}{2!} + \cdots + \frac{1}{n!} + R_n,$$

with

$$0 < \frac{3}{(n+1)!}.$$

Multiplying throughout by $n!$, we get

$$n!\frac{p}{q} = n! + n! + \frac{n!}{2!} + \ldots + \frac{n!}{n!} + n!R_n.$$

But since $n > q$, $\frac{n!}{q}$ is an integer; so is $\frac{n!}{j!}$ for any positive integer $j \leq n$. This implies that $n!R_n$ is an integer. But

$$0 < n!R_n < \frac{n!3}{(n+1)!} = \frac{3}{n+1},$$

and this gives a contradiction because $n > 3$, implying that $3/(n+1)$ is < 1.

Hence e must be irrational!

$\square$

11 Uniform convergence, Taylor series, Complex series

In this chapter we will lay down the basic results underlying the convergence – **pointwise** and **uniform** of infinite series of functions. No proofs will be given, but the student should feel free to make use of them, of course taking care to apply them only when the relevant hypotheses are satisfied. We will use this theory to study the ever important **Taylor series expansions** of various functions, including the exponential, trigonometric and simple rational functions, as well as the logarithm. We will at the end extend these functions to the domain of complex numbers to the extent possible.

11.1 Infinite series of functions, convergence

Let $\{f_n\}$ be a sequence of functions on a subset X of $\mathbb{R}$ or $\mathbb{C}$. We will say that this sequence is **pointwise convergent** with limit f on X iff for every $x \in X$, the sequence $\{f_n(x)\}$ of numbers converges to $f(x)$. In other words, for every $\varepsilon > 0$, there is a positive number $N(x)$ such that

$$n > N(x) \implies |f(x) - f_n(x)| < \varepsilon. \tag{11.1.1}$$

It is natural to wonder if $N(x)$ can be taken to be a number N, say, which is **independent of** x. In such a case, we will say that f_n is **uniformly convergent** with limit f on X.

Uniform convergence is a very important concept for Calculus. To elaborate, we have the following useful result:

Theorem 11.1.2 *Let $\{f_n\}$ be a sequence of functions on a finite interval I in $\mathbb{R}$, and let f be a function on the same interval I. Then*

*(a) (**Integration**) If $f_n \to f$ uniformly on I, and if f_n, f are integrable on I, then*

$$\int_I f(x)dx = \lim_{n\to\infty} \int_I f_n(x)dx.$$

*(b) (**Continuity**) If $f_n \to f$ uniformly on I, and if the f_n are continuous, then f is also continuous.*

*(c) (**Differentiability**) Suppose $\{f_n\}$ converge pointwise to f, where each f_n is differentiable with f_n' integrable. Further assume that the sequence $\{f_n'\}$ converges uniformly to a continuous function ϕ on I. Then f is differentiable and*

$$f'(x) = \lim_{n\to\infty} f_n'(x),$$

for all $x \in I$.

Now consider any **infinite sum of functions** on a subset X of $\mathbb{R}$ or $\mathbb{C}$, given as

$$\sum_{n=0}^{\infty} f_n. \tag{11.1.3}$$

To know if this makes sense, we need to look at the **partial sums** s_n defined by

$$s_n = \sum_{j=0}^{n} f_j. \tag{11.1.4}$$

We will say that the sum S converges uniformly, resp. **pointwise**, to a function S on X iff the sequence $\{s_n\}$ converges to S uniformly, resp. pointwise, on X.

Theorem 11.1.2 has a natural analog for infinite sums.

Theorem 11.1.5 *Suppose we are given $\sum_{n=0}^{\infty} f_n, S$ on an interval I in $\mathbb{R}$.*

*(a) (**Integrability**) If $\sum_{n=0}^{\infty}$ converges uniformly to S in I, and if the f_n, f are integrable there, then*

$$\int_I S(x)dx = \sum_{n=1}^{\infty} \int_I f_n(x)dx.$$

*(b) (**Continuity**) If $\sum_{n=0}^{\infty}$ converges uniformly to S in I, and if the f_n are continuous, the so is S.*

*(c) (**Differentiability**) Suppose $\sum_{n=1}^{\infty} f_n$ converges pointwise to S, where each f_n is differentiable with f_n' integrable. Further assume that the sum $\sum_{n=1}^{\infty} f_n'$ converges uniformly to a continuous function T on I. Then S is differentiable and*

$$S'(x) = \sum_{n=1}^{\infty} f_n'(x),$$

for all $x \in I$.

Here is a very helpful test to determine if a given infinite series converges uniformly or not.

Theorem 11.1.6 (Weierstrass's test) *Let $\{f_n\}$ be a sequence of functions on an interval I in $\mathbb{R}$, and let $\{M_n\}$ be a sequence of positive numbers such that*

(i) $|f_n(x)| \leq M_n$ for all n, and for all x; and

(ii) the infinite series of numbers $\sum_{n=1}^{\infty} M_n$ converges.

Then

(a) For all $x \in I$, the series $\sum_{n=1}^{\infty} f_n(x)$ converges absolutely; and

(b) the sum $\sum_{n=1}^{\infty} f_n$ *converges uniformly on* I *to the function*

$$S(x) = \sum_{n=1}^{\infty} f_n(x).$$

Here is an **example** illustrating the usefulness of this test. For any integer $k \geq 2$, consider the series

$$\sum_{n=1}^{\infty} \frac{\sin nx}{n^k}. \tag{11.1.7}$$

Put

$$M_n = \frac{1}{n^k}.$$

Then, since $\sin nx$ is always bounded between -1 and 1,

$$\left| \frac{\sin nx}{n^k} \right| \leq M_n \quad \forall\, n \geq 1.$$

Moreover, since $k > 1$,

$$\sum_{n=1}^{\infty} M_n$$

converges. So we may apply Weierstrass's test and conclude that the series (11.1.7) converges.

Note that if we take the above series, but with $k = 1$, then the test does not apply as the sum of $1/n$ diverges. But that does not mean that the series $\sum_{n\geq 1} \sin nx\, n^{-1}$ is divergent for all x. It only means that the method does not apply. In fact one can show, with some trouble, that this series is convergent for x in $[t, \pi - t]$, for any $t > 0$. Of course it is problematic when x is a multiple of π.

11.2 Taylor series

A series of the form

$$\sum_{n=0}^{\infty} a_n x^n \tag{11.2.1}$$

is called a **power series**.

Since the power functions $x \to x^n$ are integrable and differentiable, in fact any number of times, we may specialize Theorem 11.1.5 to the case of power series and obtain the following:

Theorem 11.2.2 For any sequence $\{a_n\}$ of real numbers, put

$$S(x) = \sum_{n=0}^{\infty} a_n x^n$$

(i) If the series is **uniformly** convergent in a closed interval $[a, b]$, then S is integrable on $[a, b]$ and

$$\int_a^b S(x)dx = \lim_{n\to\infty} a_n \frac{b^{n+1} - a^{n+1}}{n+1}.$$

(ii) Suppose the series converges pointwise, and more importantly, its series of derivatives, namely

$$T(x) = \sum_{n=1}^{\infty} n a_n x^{n-1}$$

converges **uniformly** in $[a, b]$. Then S is differentiable with derivative T.

A particular kind of power series, which is of great utility, is the *so called* **Taylor series** (at 0), which is of the form

$$\sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!} x^n \tag{11.2.3}$$

for some infinitely differentiable function f.

Theorem 11.2.2 says that if a power series is uniformly convergent we can integrate it term by term to get its integral, and if the series of derivatives is also uniformly convergent, then we can differentiate it term by term to get the derivative. So, roughly speaking, when we have uniform convergence of the appropriate power series we are allowed to integrate and differentiate term by term. What is left to do is to find a criterion to know when a power series is uniformly convergent. We also want this criterion to tell us when the series and the formally derived series both converge uniformly. Here is the completely satisfactory result giving us what we need:

Theorem 11.2.4 *Let $\{a_n\}$ be a sequence of real numbers. Suppose there is a positive real number c such that*

$$\sum_{n=0}^{\infty} a_n c^n$$

converges. Pick any positive number $b < c$. Then each of the power series

$$S(x) = \sum_{n=0}^{\infty} a_n x^n$$

and

$$T(x) = \sum_{n=1}^{\infty} n a_n x^{n-1}$$

converges absolutely and uniformly in $[-b, b]$. Moreover,

$$S'(x) = T(x) \quad \forall\, x \in (-c, c).$$

The proof of this result, which we will skip, makes clever use of the Weierstrass test and Theorem 11.2.2.

Let us try to understand the **exponential function** using this. We have seen that

$$e^x \;=\; 1 + x + \frac{x^2}{2!} + \ldots + \frac{x^n}{n!} + R_n(e^x; 0),$$

with

$$|R_n(e^x; 0)| \;\le\; e^x \frac{|x|^{n+1}}{(n+1)!}.$$

Consequently, for any positive c, the remainder term $R_n(e^c; 0)$ goes to zero as n goes to infinity. This gives us the convergent infinite series expression

$$e^c \;=\; 1 + c + \frac{c^2}{2!} + \frac{c^3}{3!} + \ldots$$

Applying Theorem 11.2.4 we see then that the **Taylor series**

$$e^x \;=\; 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \ldots \tag{11.2.5}$$

converges absolutely and uniformly on $[-c, c]$ for any $c > 0$. Moreover, its derivative is given by differentiating the series term by term, which gives back e^x, as expected.

By a similar argument, we also get the Taylor series expressions for the sine and cosine functions, valid for all x:

$$\sin x \;=\; x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \ldots \tag{11.2.5 $-$ i}$$

$$\cos x \;=\; 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \ldots \tag{11.2.5 $-$ ii}$$

We can also apply Theorem 11.2.4 and differentiate the expression for $\sin x$ term by term, and as expected, one gets the series expression for $\cos x$.

The series

$$\arctan x \;=\; x - \frac{x^3}{3} + \frac{x^5}{5} - \ldots \tag{11.2.7}$$

is, as we saw before, convergent at $x = 1$. Thus by Theorem 11.2.4, it converges absolutely in $-1 < x < 1$ and uniformly in $[-b, b]$ for any positive $b < 1$.

We claim that this Taylor series for $\arctan x$ does not converge at any $x > 1$. Indeed, if it did, then by Theorem 11.2.4, it would converge absolutely at $x = 1$ which it does not. So the number 1, which some call the *radius of convergence* of the Taylor series (11.2.7), is the boundary beyond which there is no convergence.

Taking the derivative of the series (11.2.7) term by term, and applying Theorem 11.2..4, we get (as expected)

$$\frac{1}{1+x^2} = 1 - x^2 + x^4 - x^6 + \dots, \tag{11.2.8}$$

which is valid in $\{|x| < 1\}$. This certainly does not converge at $x = -1$. Even at $x = 1$ the series makes no sense, but the left hand side makes sense there and equals $\frac{1}{2}$.

Incidentally, the Taylor series for $\arctan x$ converges at $x = -1$, with

$$\arctan(-1) = -\arctan 1 = -\frac{\pi}{4}.$$

It should be noted, despite what the arctan function may suggest, that in Theorem 11.2.4, the absolute convergence of $S(x)$ is asserted only for $|x| < c$, and one can say nothing in general about the convergence, or the lack of it, at $x = -c$. Sometimes it does happen that $S(x)$ is divergent at $-c$. The simplest example illustrating this is the Taylor series for $\log(1+x)$ at $x = 0$, which we claim to be

$$\log(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \dots \tag{11.2.9}$$

absolutely convergent in $-1 < x < 1$ and uniformly convergent in $[-b, b]$ for any positive $b < 1$. One way to see this is the following:

Since $\log(1+x)$ is (defined and) differentiable in $x > -1$ with derivative $\frac{1}{1+x}$, we could find the Taylor series expansion of $\frac{1}{1+x}$ first and then differentiate.

The identity (11.2.8) implies that

$$\frac{1}{1+x} = 1 - x + x^3 - x^4 + \dots \tag{11.2.10}$$

which is absolutely convergent in $0 \leq x < 1$. Then it is also absolutely convergent in $(-1, 1)$. Pick any positive $b < 1$. Then since the series (11.2.10) converges at c for any c with $b < c < 1$. So we may apply Theorem 11.2.4 and conclude that (11.2.10) converges uniformly in $[-b, b]$. We can also differentiate term by term, by this Theorem, and obtain the desired expansion (11.2.9), which is valid in $-1 \leq x \leq 1$.

At $x = 1$, the series (11.2.9) is alternating and converges to $\log 2$. But it is important to note, however, that it is divergent at $x = -1$.

A consequence of (11.2.9) is the series for $\log(1-x)$, given by

$$-\log(1-x) = \log\left(\frac{1}{1-x}\right) = x + \frac{x^2}{2} + \frac{x^3}{3} + \dots \tag{11.2.11}$$

A very useful thing to remember, when computing the Taylor series of various functions is the following

Lemma 11.2.12 *For $j = 1, 2$, consider power series*

$$S_j(x) \;=\; \sum_{n=0}^{\infty} a_{n,j} x^n$$

If they are both uniformly convergent in some $[a, b]$, then for all scalars α, β, the power series

$$\sum_{n=0}^{\infty} (\alpha a_{n,1} + \beta a_{n,2}) x^n$$

converges uniformly in $[a, b]$ to $\alpha S_1(x) + \beta S_2(x)$.

In particular, the sum or difference of two uniformly convergent power series is again uniformly convergent.

11.3 Complex power series

It is very important to understand the power series

$$\sum_{n=0}^{\infty} a_n z^n \tag{11.3.1}$$

where the coefficients a_n are complex numbers and z is in $\mathbb{C}$.

We will say that such a series is absolutely convergent if the associated (non-negative) real series

$$\sum_{n=0}^{\infty} |a_n| \cdot |z|^n \tag{11.3.2}$$

converges. For example, the **complex exponential function**

$$e^z \;=\; 1 + z + \frac{z^2}{2!} + \frac{z^3}{3!} + \dots \tag{11.3.3}$$

is absolutely convergent for all z because the corresponding series of absolute values, namely

$$1 + |z| + \frac{|z|^2}{2!} + \frac{|z|^3}{3!} + \dots$$

is convergent, it being the Taylor series of the real exponential $e^{|z|}$.

Here is the main result of this section.

Theorem 11.3.4 *Suppose the complex power series*

$$S(z) \;=\; \sum_{n=0}^{\infty} a_n z^n$$

converges at some non-zero complex number z_0. Then either this series is absolutely convergent for all z in $\mathbb{C}$, or there is a positive real number $R > 0$ such that

$S(z)$ is absolutely convergent in $\{|z| < R\}$; and

$S(z)$ diverges for any z with $|z| > R$.

Moreover, $S(z)$ is differentiable at any z with $|z| < R$ with derivative

$$S'(z) \;=\; \sum_{n=1}^{\infty} n a_n z^{n-1}.$$

R is called the **radius of convergence** of the power series. Note that the Theorem says nothing about the convergence at the points z with $|z| = 1$. Some people would say that the radius of convergence is infinite when $S(z)$ is absolutely convergent everywhere in $\mathbb{C}$.

One can define the complex extensions of the various functions we have encountered like the sine and cosine functions in terms of their power series. To be specific we have

$$\sin z \;=\; z - \frac{z^3}{3!} + \frac{z^5}{5!} - \frac{z^7}{7!} + \ldots \tag{11.3.5}$$

$$\cos z \;=\; 1 - \frac{z^2}{2!} + \frac{z^4}{4!} - \frac{z^6}{6!} + \ldots \tag{11.3.6}$$

Both these power series have infinite radii of convergence.

It is crucial to note, by comparing the power series expansions, that

$$e^{iz} \;=\; \cos z + i \sin z. \tag{11.3.7}$$

This leads to the identities

$$\sin z \;=\; \frac{e^{iz} - e^{-iz}}{2i} \tag{11.3.8}$$

and

$$\cos z \;=\; \frac{e^{iz} + e^{-iz}}{2}. \tag{11.3.9}$$

We will be remiss if we forget the **complex logarithm**

$$\log(1+z) \;=\; z - \frac{z^2}{2} + \frac{z^3}{3} - \frac{z^4}{4} + \ldots \tag{11.3.10}$$

which is absolutely convergent in $\{|z| < 1\}$. We have already seen that it converges at $z = 1$ and diverges at $z = -1$.

What we now have a is a very interesting situation for $\log(1+z)$. On the one hand, it makes sense at any complex number z with $|z| < 1$. On the other, it makes sense whenever

$1+z$ is a positive real number, though it is not given by its Taylor series (11.3.10) if $|z|>1$. Thus $\log(1+z)$ is defined on the following strange-shaped subset of $\mathbb{C}$:

$$\{z \in \mathbb{C} \,|\, |z|<1 \quad \text{or} \quad z \in (-1, \infty) \subset \mathbb{R}\}. \tag{11.3.11}$$

The natural question which arises immediately is whether this function can be given sense at all points of $\mathbb{C}$, or at least on a larger subset. The quick answer is that it can be defined on the *complement* in $\mathbb{C}$ of the half-line $(-\infty, -1] \subset \mathbb{R}$. On all of $\mathbb{C}$ it can only be defined as a multi-valued function. To understand all these intriguing matters better, one will need to understand a fundamental mathematical concept called *analytic continuation*, and for that one will need to take a course in *Complex Variables*.